
JMIR Human Factors

Impact Factor (2023): 2.6

Volume 9 (2022), Issue 2 ISSN 2292-9495 Editor in Chief: Andre Kushniruk, BA, MSc, PhD, FACMI

Contents

Original Papers

- Persuasive Design Techniques and App Design Recommendations to Improve Health Workforce Capability in Rural Health Professionals: What Do Users Want and How Does an App Help? ([e35094](#))
Sabrina Pit, Robyn Ramsden, Aaron Tan, Kristy Payne, James Barr, Benjamin Eames, Mike Edwards, Richard Colbran. 6
- Exploring Human-Data Interaction in Clinical Decision-making Using Scenarios: Co-design Study ([e32456](#))
Helena Tendedez, Maria-Angela Ferrario, Roisin McNaney, Adrian Gradinar. 22
- Combining Persuasive System Design Principles and Behavior Change Techniques in Digital Interventions Supporting Long-term Weight Loss Maintenance: Design and Development of eCHANGE ([e37372](#))
Rikke Asbjørnsen, Jøran Hjelmæsæth, Mirjam Smedsrød, Jobke Wentzel, Marianne Ollivier, Matthew Clark, Julia van Gemert-Pijnen, Lise Solberg Nes. 43
- Assessing the Usability of a Clinical Decision Support System: Heuristic Evaluation ([e31758](#))
Hwayoung Cho, Gail Keenan, Olatunde Madandola, Fabiana Dos Santos, Tamara Macieira, Ragnhildur Bjarnadottir, Karen Priola, Karen Dunn Lopez. 189
- Adoption of a Postoperative Pain Self-Report Tool: Qualitative Study ([e33706](#))
Bram Thiel, Inez Iao, Joris Smid, Emmy de Wit, Seppe Koopman, Bart Geerts, Marc Godfried, Cor Kalkman. 202
- A Mobile App for Advance Care Planning and Advance Directives (Accordons-nous): Development and Usability Study ([e34626](#))
Céline Schöpfer, Frederic Ehrlé, Antoine Berger, Catherine Bollondi Pauly, Laurence Buytaert, Camille De La Serna, Florence Hartheiser, Thomas Fassier, Christine Clavien. 226
- Understanding the Potential of Mental Health Apps to Address Mental Health Needs of the Deaf and Hard of Hearing Community: Mixed Methods Study ([e35641](#))
Judith Borghouts, Martha Neary, Kristina Palomares, Cinthia De Leon, Stephen Schueller, Margaret Schneider, Nicole Stadnick, Dana Mukamel, Dara Sorkin, Dakota Brown, Shannon McCleerey-Hooper, Gloria Moriarty, Elizabeth Eikey. 245
- Process and Information Needs When Searching for and Selecting Apps for Smoking Cessation: Qualitative Study Using Contextual Inquiry ([e32628](#))
Ylva Hendriks, Sebastiaan Peek, Maurits Kaptein, Inge Bongers. 256
- Factors Influencing Clinician Trust in Predictive Clinical Decision Support Systems for In-Hospital Deterioration: Qualitative Descriptive Study ([e33960](#))
Jessica Schwartz, Maureen George, Sarah Rossetti, Patricia Dykes, Simon Minshall, Eugene Lucas, Kenrick Cato. 279

A Web-Based Stratified Stepped Care Mental Health Platform (TourHeart): Semistructured Interviews With Stakeholders (e35057)	
Emily Tsoi, Winnie Mak, Connie Ho, Gladys Yeung.	294
#BingeDrinking—Using Social Media to Understand College Binge Drinking: Qualitative Study (e36239)	
Madison Cirillo, Jennifer Halbert, Jessica Smith, Nour Alamiri, Karen Ingersoll.	308
Features and Components Preferred by Adolescents in Smartphone Apps for the Promotion of Physical Activity: Focus Group Study (e33972)	
Alex Domin, Yacine Ouzzahra, Claus Vögele.	320
Use of Health Care Chatbots Among Young People in China During the Omicron Wave of COVID-19: Evaluation of the User Experience of and Satisfaction With the Technology (e36831)	
Yi Shan, Meng Ji, Wenxiu Xie, Xiaomin Zhang, Xiaobo Qian, Rongying Li, Tianyong Hao.	330
Developing mHealth to the Context and Valuation of Injured Patients and Professionals in Hospital Trauma Care: Qualitative and Quantitative Formative Evaluations (e35342)	
Thymen Houwen, Miel Vugts, Koen Lansink, Hilco Theeuwes, Nicky Neequaye, M Beerekamp, Margot Joosen, Mariska de Jongh.	346
Implementation and Evaluation of COVIDCare@Home, a Family Medicine–Led Remote Monitoring Program for Patients With COVID-19: Multimethod Cross-sectional Study (e35091)	
Celia Laur, Payal Agarwal, Kelly Thai, Vanessa Kishimoto, Shawna Kelly, Kyle Liang, R Bhatia, Onil Bhattacharyya, Danielle Martin, Geetha Mukerji.	363
Caregiver Expectations of Interfacing With Voice Assistants to Support Complex Home Care: Mixed Methods Study (e37688)	
Ryan Tennant, Sana Allana, Kate Mercer, Catherine Burns.	377
Interdisciplinary Collaborations in Digital Health Research: Mixed Methods Case Study (e36579)	
Grit Krause-Jüttler, Jürgen Weitz, Ulrich Bork.	394
Proposal for Post Hoc Quality Control in Instrumented Motion Analysis Using Markerless Motion Capture: Development and Usability Study (e26825)	
Hanna Röhlng, Patrik Althoff, Radina Arsenova, Daniel Drebing, Norman Gigengack, Anna Chorschew, Daniel Kroneberg, Maria Rönnefarth, Tobias Ellermeyer, Sina Rosenkranz, Christoph Heesen, Behnouth Behnia, Shigeki Hirano, Satoshi Kuwabara, Friedemann Paul, Alexander Brandt, Tanja Schmitz-Hübsch.	412
Evaluation of the First Year(s) of Physicians Collaboration on an Interdisciplinary Electronic Consultation Platform in the Netherlands: Mixed Methods Observational Study (e33630)	
Sanne Sanavro, Henk van der Worp, Danielle Jansen, Paul Koning, Marco Blanker, Prisma Platform Study Group.	426
Implementation of a Web-Based Tool for Shared Decision-making in Lung Cancer Screening: Mixed Methods Quality Improvement Evaluation (e32399)	
Julie Lowery, Angela Fagerlin, Angela Larkin, Renda Wiener, Sarah Skurla, Tanner Caverly.	435
Requirements for a Bespoke Intensive Care Unit Dashboard in Response to the COVID-19 Pandemic: Semistructured Interview Study (e30523)	
Brittany Davidson, Katuska Ferrer Portillo, Marceli Wac, Chris McWilliams, Christopher Bourdeaux, Ian Craddock.	448
Enhancing Clinical Information Display to Improve Patient Encounters: Human-Centered Design and Evaluation of the Parkinson Disease-BRIDGE Platform (e33967)	
Ethan Brown, Erica Schleimer, Ian Bledsoe, William Rowles, Nicolette Miller, Stephan Sanders, Katherine Rankin, Jill Ostrem, Caroline Tanner, Riley Bove.	462

An Electronic Data Capture Tool for Data Collection During Public Health Emergencies: Development and Usability Study (e35032)	
Joan Brown, Manas Bhatnagar, Hugh Gordon, Jared Goodner, J Cobb, Karen Lutrick.	472
Usability of the Swedish Accessible Electronic Health Record: Qualitative Survey Study (e37192)	
Maria Hägglund, Isabella Scandurra.	482
The Sociodemographic Digital Divide in Mobile Health App Use Among Clients at Outpatient Departments in Inner Mongolia, China: Cross-sectional Survey Study (e36962)	
Li Cao, Virasakdi Chongsuvivatwong, Edward McNeil.	494
Understanding People With Chronic Pain Who Use a Cognitive Behavioral Therapy–Based Artificial Intelligence Mental Health App (Wysa): Mixed Methods Retrospective Observational Study (e35671)	
Saha Meheli, Chaitali Sinha, Madhura Kadaba.	508
Identifying Contextual Factors and Strategies for Practice Facilitation in Primary Care Quality Improvement Using an Informatics-Driven Model: Framework Development and Mixed Methods Case Study (e32174)	
Jiancheng Ye, Donna Woods, Jennifer Bannon, Lucy Bilaver, Gayle Kricke, Megan McHugh, Abel Kho, Theresa Walunas.	519
A Participatory Design Approach to Develop Visualization of Wearable Actigraphy Data for Health Care Professionals: Case Study in Qatar (e25880)	
Kamran Khowaja, Wafa Syed, Meghna Singh, Shahrad Taheri, Odette Chagoury, Dena Al-Thani, Michaël Aupetit.	534
A Nurse-Led Multimedia Intervention to Increase Patient Participation in Recovery After Knee Arthroplasty: Hybrid Type II Implementation Study (e36959)	
Jo McDonall, Bernice Redley, Patricia Livingston, Ana Hutchinson, Richard de Steiger, Mari Botti.	557
Co-creation of a Serious Game About Radiotherapy: Participatory Action Research Study With Children Treated for Cancer (e34476)	
Catarina Cederved, Jon Back, Charlotte Ångström-Brännström, Gustaf Ljungman, Gunn Engvall.	570
Development of a Digital Tool for People With a Long-Term Condition Using Stroke as a Case Example: Participatory Design Approach (e35478)	
Emma Kjörk, Katharina Sunnerhagen, Åsa Lundgren-Nilsson, Anders Andersson, Gunnel Carlsson.	582
Smartphone Alcohol Use Disorder Recovery Apps: Cross-sectional Survey of Behavioral Intention to Use (e33493)	
Rijuta Menon, Julien Meyer, Pria Nippak, Housne Begum.	599
Evaluating User Feedback for an Artificial Intelligence–Enabled, Cognitive Behavioral Therapy–Based Mental Health App (Wysa): Qualitative Thematic Analysis (e35668)	
Tanya Malik, Adrian Ambrose, Chaitali Sinha.	608
Designing Tangibles to Support Emotion Logging for Older Adults: Development and Usability Study (e34606)	
Daniel Gooch, Vikram Mehta, Avelie Stuart, Dmitri Katz, Mohamed Bennasar, Mark Levine, Aroscha Bandara, Bashar Nuseibeh, Amel Bennaceur, Blaine Price.	618
Determinants of Laypersons’ Trust in Medical Decision Aids: Randomized Controlled Trial (e35219)	
Marvin Kopka, Malte Schmieiding, Tobias Rieger, Eileen Roesler, Felix Balzer, Markus Feufel.	637
A Mobile App for Prevention of Cardiovascular Disease and Type 2 Diabetes Mellitus: Development and Usability Study (e35065)	
Vera Buss, Marlien Varnfield, Mark Harris, Margo Barr.	653

Audio Recording Patient-Nurse Verbal Communications in Home Health Care Settings: Pilot Feasibility and Usability Study (e35325)
 Maryam Zolnoori, Sasha Vergez, Zoran Kostic, Siddhartha Jonnalagadda, Margaret V McDonald, Kathryn Bowles, Maxim Topaz. 673

Application of a Web-based Self-assessment Triage Tool During the COVID-19 Pandemic: Descriptive Study (e34134)
 Anna Nowicka, Jakub Jaszczak, Anna Szymanek Pasternak, Krzysztof Simon. 688

Patient Portal Messaging for Asynchronous Virtual Care During the COVID-19 Pandemic: Retrospective Analysis (e35187)
 Ming Huang, Aditya Khurana, George Mastorakos, Andrew Wen, Huan He, Liwei Wang, Sijia Liu, Yanshan Wang, Nansu Zong, Julie Prigge, Brian Costello, Nilay Shah, Henry Ting, Jungwei Fan, Christi Patten, Hongfang Liu. 701

Modeling Trust in COVID-19 Contact-Tracing Apps Using the Human-Computer Trust Scale: Online Survey Study (e33951)
 Sonia Sousa, Tiina Kalju. 711

Health Professionals' eHealth Literacy and System Experience Before and 3 Months After the Implementation of an Electronic Health Record System: Longitudinal Study (e29780)
 Lars Kayser, Astrid Karnoe, Emily Duminski, Svend Jakobsen, Rikke Terp, Susanne Dansholm, Michael Roeder, Gustav From. 734

A Web-Based Health Application to Translate Nutrition Therapy for Cardiovascular Risk Reduction in Primary Care (PortfolioDiet.app): Quality Improvement and Usability Testing Study (e34704)
 Meaghan Kavanagh, Laura Chiavaroli, Andrea Glenn, Genevieve Heijmans, Shannan Grant, Chi-Ming Chow, Robert Josse, Vasanti Malik, William Watson, Aisha Lofters, Candice Holmes, Julia Rackal, Kristie Srichaikul, Diana Sherifali, Erna Snelgrove-Clarke, Jacob Udell, Peter Juni, Gillian Booth, Michael Farkouh, Lawrence Leiter, Cyril Kendall, David Jenkins, John Sievenpiper. 748

Effective Communication of Personalized Risks and Patient Preferences During Surgical Informed Consent Using Data Visualization: Qualitative Semistructured Interview Study With Patients After Surgery (e29118)
 Undina Gisladdottir, Drashko Nakikj, Rashi Jhunjhunwala, Jasmine Panton, Gabriel Brat, Nils Gehlenborg. 761

A Reference Architecture for Data-Driven and Adaptive Internet-Delivered Psychological Treatment Systems: Software Architecture Development and Validation Study (e31029)
 Suresh Mukhiya, Yngve Lamo, Fazle Rabbi. 776

Reviews

Chinese Americans' Use of Patient Portal Systems: Scoping Review (e27924)
 Katharine Lawrence, Stella Chong, Holly Krelle, Timothy Roberts, Lorna Thorpe, Chau Trinh-Shevrin, Stella Yi, Simona Kwon. 71

Functional and Technical Aspects of Self-management mHealth Apps: Systematic App Search and Literature Review (e29767)
 Lyan Alwakeel, Kevin Lano. 88

Assessing the Legal Aspects of Information Security Requirements for Health Care in 3 Countries: Scoping Review and Framework Development (e30050)
 Prosper Yeng, Muhammad Fauzi, Luyi Sun, Bian Yang. 114

Digital Health Opportunities to Improve Primary Health Care in the Context of COVID-19: Scoping Review (e35380)
 Cícera Silva, Rayssa Lopes, Osvaldo de Goes Bay Jr, Claudia Martiniano, Miguel Fuentealba-Torres, Ricardo Arcêncio, Luís Lapão, Sonia Dias, Severina Uchoa. 139

The Impact of Perioperative Remote Patient Monitoring on Clinical Staff Workflows: Scoping Review (e37204)	
Maria León, Valeria Pannunzio, Maaïke Kleinsmann.	158
User Perceptions of eHealth and mHealth Services Promoting Physical Activity and Healthy Diets: Systematic Review (e34278)	
Julia Bergevi, Susanne Andermo, Yohannes Woldamanuel, Unn-Britt Johansson, Maria Hagströmer, Jenny Rossen.	177
Self-care Behaviors and Technology Used During COVID-19: Systematic Review (e35173)	
Fareeya Sakur, Kanasha Ward, Neha Khatri, Annie Lau.	721

Viewpoint

Toward an Ecologically Valid Conceptual Framework for the Use of Artificial Intelligence in Clinical Settings: Need for Systems Thinking, Accountability, Decision-making, Trust, and Patient Safety Considerations in Safeguarding the Technology and Clinicians (e35421)	
Avishek Choudhury.	211

Corrigenda and Addenda

Correction: Improving Pelvic Floor Muscle Training Adherence Among Pregnant Women: Validation Study (e38175)	
Aida Jaffar, Sherina Mohd-Sidik, Chai Foo, Novia Admodisastro, Sobihatun Abdul Salam, Noor Ismail.	597

Letter to the Editor

Should Expert Surgeon Guidance Be Given Remotely? Comment on “User Experience in Remote Surgical Consultation: Survey Study of User Acceptance and Satisfaction in Real-Time Use of a Telemedicine Service” (e36681)	
Daniella Soussi, Chiara Vedi.	686

Original Paper

Persuasive Design Techniques and App Design Recommendations to Improve Health Workforce Capability in Rural Health Professionals: What Do Users Want and How Does an App Help?

Sabrina Pit^{1,2,3}, MSc, PhD; Robyn Ramsden¹, BEd, PhD; Aaron JH Tan¹, BMed, MD; Kristy Payne¹, BPsych, MCLinEpi; James Barr¹, BVisualComDesign; Benjamin Eames¹, BBus; Mike Edwards¹, BCompSc; Richard Colbran¹, BA

¹New South Wales Rural Doctors Network, Hamilton, Australia

²School of Medicine, Western Sydney University, Campbelltown, Australia

³University Centre for Rural Health, Faculty of Medicine and Health Sciences, University of Sydney, Lismore, Australia

Corresponding Author:

Sabrina Pit, MSc, PhD

New South Wales Rural Doctors Network

1/53 Cleary Street

Hamilton, 2308

Australia

Phone: 61 429455720

Email: sabrina.pit@sydney.edu.au

Abstract

Background: Health professionals' perceptions of persuasive design techniques for use in technological solutions to improve health workforce capability have not been previously explored.

Objective: This study aims to explore rural health professionals' perceptions of health workforce capability and persuasive design techniques; and translate these into recommendations for designing a health workforce capability app to increase their impact and usefulness.

Methods: Qualitative interviews with 13 rural health professionals were conducted. Subsequently, 32 persuasive techniques were used as a framework to deductively analyze the data. Persuasive design technique domains were Primary Task Support, Dialog Support, System Credibility Support, Social Support, and Cialdini's Principles of Persuasion.

Results: Persuasive design techniques can be applied across the factors that influence health workforce capability including health and personal qualities; competencies and skills; values, attitudes, and motivation; and factors that operate outside of work and at the team, organizational, and labor market levels. The majority of the 32 persuasive design techniques were reflected in the data from the interviews and led to recommendations as to how these could be translated into practice, with the exception of scarcity. Many suggestions and persuasive design techniques link back to the need for tailored and localized solutions such as the need for country-specific-based evidence, the wish for localized communities of practice, learning from other rural health professionals, and referral pathways to other clinicians. Participants identified how persuasive design techniques can optimize the user experience to help meet rural health professionals needs for more efficient systems to improve patient access to care, quality care, and to enable working in interprofessional team-based care. Social inclusion plays a vital role for health professionals, indicating the importance of the Social Support domain of persuasive techniques. Overall, health professionals were open to self-monitoring of their work performance and some professionals used wearables to monitor their health.

Conclusions: Rural health professionals' perceptions of health workforce capability informed which persuasive design techniques can be used to optimize the user experience of an app. These were translated into recommendations for designing a health workforce capability app to increase likelihood of adoption. This study has also contributed to the further validation of the Persuasive Systems Design model through empirically aligning elements of the model to increase persuasive system content and functionality with real-world applied data, in this case the health workforce capability of rural health professionals. Our findings confirm that these techniques can be used to develop a future prototype of an app that may assist health professionals in improving or maintaining their health workforce capability which in turn may increase recruitment and retention in rural areas. Health professionals need

to be central during the design phase. Interventions are needed to provide a supportive environment to rural and remote health professionals to increase their rural health workforce capability.

(*JMIR Hum Factors* 2022;9(2):e35094) doi:[10.2196/35094](https://doi.org/10.2196/35094)

KEYWORDS

health; wellness; mobile apps; persuasive strategies; behavior change; review; health workforce; capability; career; employment; rural; workforce planning; mHealth; mobile health; digital health; health professional; user experience; health application; task support; social support; dialog support

Introduction

Background

Globally, people in rural areas are disadvantaged when it comes to seeking health care [1,2]. This manifests as poorer health outcomes and has, in part, been directly attributed to persistent recruitment and retention challenges in rural areas [3]. Additionally, high levels of burnout are reported among rural health professionals across the globe [4,5]. For example, 47.6% of rural Chinese primary health care workers reported moderate burnout and 3% severe burnout [5].

Health Workforce Capability

There is a growing body of evidence that aims to address the barriers to, and enablers of, recruiting and retaining a rural health workforce [3,6]. For example, multiple initiatives are aimed at incentivizing work in rural areas for health professionals [7], and others demonstrate the potential effectiveness of regulatory change such as increased scope of practice for nurse practitioners [8]. Specifically, workforce capability plays a significant role in the attraction, recruitment, and maintenance of health practitioners in a rural setting [9]. Health workforce capability describes “a health professional’s overall ability to fulfil their health care role” [9]. Health workforce capability can be defined as “the intersection between individual capacity and ability to adapt to work considering the whole of health care context, including the labor market, population needs, family, schools, partner, education, and social options.” Thus, health workforce capability is a complex construct that consists of much more than just clinical competence. It is a holistic concept that considers the multidirectional interactions both internal and external to the individual practitioner [9-11]. Internally, it considers interactions between an individual practitioner’s personal and professional spheres, both of which are complicated domains that vary greatly between individuals. Externally, workforce capability takes into account the interactions between the professional and their environment, including their employer, co-workers, and social circles. For rural health professionals, it also takes into account patients, clients, and the communities they serve. Despite the clear importance of understanding capability in the context of the rural health workforce, it is an underexplored area.

Digital Technology and Health Workforce Capability

Before the role of technology in the clinical context was forced into the limelight due to COVID-19 [12], rural health professionals globally had long been using digital solutions [12] to bridge large geographical distances. Complementary to this, recent studies have demonstrated technology’s role in supporting

the workforce capability of the rural health workforce [13]. Indeed, it has been suggested that employing technology solutions in the rural context can improve the capability and retention of rural health professionals [13]. What then are the appropriate solutions to improve health workforce capability? To answer this question, the authors of this paper undertook 2 previous studies [10,11].

The first study investigated the wants and needs of the rural health workforce to improve their capability and the potential role of technology in assisting them [10]. Theories around the acceptability of information technology solutions, including the Mobile Application Rating Scale (MARS) [14], the Technology Acceptance Model (TAM) [15], and the Health Information Technology Acceptance Model (HITAM) [16] were used to address this question. The study identified 7 factors that had the greatest influence on health workforce capability: health and personal qualities; competencies and skills; values, attitudes, and motivation; and factors that operate outside of work and at the team, organizational, and labor market levels.

The second study consisted of reviewing apps in Google Play Store and technical design elements that allow the technology solution to fulfill the wants and needs of health professionals [11]. Persuasive design techniques can be used to encourage people to use a product or to take certain actions or to make certain positive decisions. Persuasion is a complex concept and can use computer-human or computer-mediated approaches. Computer-human approaches are computer driven, yet need to be programmed, so the developer has a critical role to play and has a large impact on design. Computer-mediated persuasion [17] occurs through using digital social communication tools. In the health context, this could include, for example, chats, interactive webinars, and virtual clinics where health professionals would persuade each other. Specifically, this second study investigated persuasive design techniques used to improve use of existing capability-building-related apps, and offered some basic suggestions for incorporating these techniques into a health workforce capability app to increase its persuasiveness. It used theories on persuasive design elements that influence behavioral change, including the Persuasive System Design Model (PSD Model) [17] and Cialdini’s Principles of Persuasion [18,19]. A total of 32 persuasive design techniques were studied clustered by 5 design features categories: Primary Task Support, Dialog Support, System Credibility Support, Social Support, and Cialdini’s Principles. The design techniques are further explained in Tables 1-5. While several studies have tried to understand the 2 questions independently [20,21], to our knowledge, none have examined improving health workforce capability on their own, let alone

mapping the 2 together. To fill this gap, this study will map persuasive design techniques with the health workforce capability needs and perspectives of rural health professionals within the context of improving health workforce capability.

Table 1. Persuasive design techniques—definitions and implementation examples for developing a health workforce capability app (Domain: Primary Task Support).

Primary task support	Persuasive design technique definition ^a	Quote qualitative interviews health professionals	Implementation examples
Reduction	A system that reduces complex behavior into simple tasks helps users perform the target behavior, and it may increase the benefit-to-cost ratio of a behavior.	<ul style="list-style-type: none"> • <i>Because there's no complications of magic <online course>. I just put in my password and that of course, comes up. It comes up where I left off. It tells me what I've done. It goes through everything logically. It has videos and writings and readings and downloads. And it's all so simple. [ID number 13, female, GP^b]</i> • <i>So maybe an app or any support to give you this feeling, you are confident, you didn't leave anything behind. Maybe to add the patient's data or - Patient X, for example, had this and that, and he needs to be checked before this date. [ID number 2, male, GP]</i> • <i>There's already heaps of computer stuff, Best Practice had a lot of various functionalities and so forth already embedded in it, for instance, the system for reminding patients' appointments. That's not the difficulty. The difficulty is more on - for us anyway - is more on the side of if someone needs a physio or that sort of - it's probably case management is the thing that helps a lot, having that sort of thing and having someone refer to an allied health person easily so they can be dealt with easy, that's where to my mind, that holistic approach, that's, to me, the capabilities thing I guess. [ID number 14, male, GP]</i> 	<ul style="list-style-type: none"> • Predetermined pathways to work on health workforce capability. • List of useful services. • GP can book an appointment for an allied health professional with an online booking system so the patient does not have to do that, which ensures the patient is booked in and holistic care is provided.
Tunneling	Using the system to guide users through a process or experience provides opportunities to persuade along the way.	<ul style="list-style-type: none"> • <i>But there are workouts that are categorised by all kinds of things, by duration, by difficulty, by area that it targets, that kind of thing, and it's also got - if you're a paying member you've also got different plans. So one of them might be a two-week plan you can follow or a six-week plan that you can follow, and you can schedule them in your calendar and they'll send alerts saying, "Your workout's due to start in 12 hours. [ID number 10, female, speech pathologist]</i> • <i>Some of those, particularly lipid screening, that that throw out a whole range of different investigations and numbers. I'm still trying to work my way through and I don't have time, although it's not high on my to do list, to sit down and do a template to work out the flags that I have all these things. It'd be great if there was already a resource to say, here it is. If it's just two points out, tell them investigate whether they're dehydrated. If it's five points out, you need to get them to the hospital, I have my own clinical judgement that I can also use to say you need to get to the hospital. Particularly since these clinics are nurse led clinics, the buck stops with me, I'm very keen to pass the buck on if I am unsure. So, it's something that I'm having to work out by myself. As I say that, another clinician showed me a website that I can go to that's a pathology website that I need to go through and pick out what I need. I don't always have web access, either. [ID number 18, female, remote nurse]</i> 	<ul style="list-style-type: none"> • An app that guides the health professional (eg, a remote nurse) through a health care protocol.
Tailoring	Information provided by the system will be more persuasive if it is tailored to the potential feeds, interests, personality, usage context, or other factors relevant to a user group.	<ul style="list-style-type: none"> • <i>It's got to be very carefully tailored. [ID number 7, male, specialist]</i> • <i>Something that would update me with some areas of interest maybe, related journal stuff, so you tick literacy or whatnot and some related journal articles would pop up. [ID number 10, female, speech pathologist]</i> 	<ul style="list-style-type: none"> • Journal articles are tailored to health professional's selected needs.
Personalization	A system that offers personalized content or services has a greater capability for persuasion.	<ul style="list-style-type: none"> • <i>So having either an unlimited time or a flexible amount of time to complete it in is great. [ID number 15, male, GP]</i> • <i>That would be something that someone else might say, "Oh, I'd like to opt out of that. Once a week is fine. [ID number 15, male, GP]</i> 	<ul style="list-style-type: none"> • Having a system that provides personalized offers to work on health workforce capability such as: <ul style="list-style-type: none"> • Australian guidelines • links to an online community of practice • Opt-in approaches.

Primary task support	Persuasive design technique definition ^a	Quote qualitative interviews health professionals	Implementation examples
Self-monitoring	A system that keeps track of one's own performance or status supports the user in achieving goals.	<ul style="list-style-type: none"> <i>You get burnout from being overworked and undervalued. And you can track that. You can track the being overworked. And you can track the value in the sense of outcomes.</i> [ID number 17, female, occupational therapist] 	<ul style="list-style-type: none"> Graphs or trends and analyses showing self-rated health workforce capability level over time and time of day.
Simulation	Systems that provide simulations can persuade by enabling users to observe immediately the link between cause and effect.	<ul style="list-style-type: none"> <i>That's very valuable to me, to see how other people actually put it into practice. So videos are quite good for that, those kind of demonstration-based things. But yeah, some summaries of research in terms of articles is also helpful.</i> [ID number 10, female speech pathologist] 	<ul style="list-style-type: none"> Videos of health professionals performing certain procedures and impact on patient outcomes.
Rehearsal	A system providing means with which to rehearse a behavior can enable people to change their attitudes or behavior in the real world.	<ul style="list-style-type: none"> <i>There's been a lot of initiative from other clinicians in my region who have wanted to use telehealth or technology for training and also professional support in regards to home visits, so things outside of the hospital environment when you're geographically isolated.</i> [ID number 17, female, occupational therapist] 	<ul style="list-style-type: none"> Simulation courses are listed and accessible to health professionals to enable rehearsal of real-world practice.

^aSource: Oinas-Kukkonen and Harjumaa Marja [17].

^bGP: general practitioner.

Table 2. Persuasive design techniques—definitions and implementation examples for developing a health workforce capability app. Domain: Dialog Support.

Dialog support	Persuasive design technique definition ^a	Quote qualitative interviews health professionals	Implementation examples
Praise	By offering praise, a system can make users more open to persuasion.	<ul style="list-style-type: none"> <i>You've got to...encourage good behaviours and not bad behaviours.</i> [ID number 7, male, pain specialist] 	<ul style="list-style-type: none"> A system can send a positive image and message when health professionals reach part of their own set goal.
Rewards	Systems that reward target behaviors may have great persuasive powers.	<ul style="list-style-type: none"> <i>It's got to be interactive, yeah, or else it's just – you've got to give people a little encouragement, yeah.</i> [ID number 7, male, pain specialist] 	<ul style="list-style-type: none"> Health care professionals get Continuing Professional Development points for completing a medical case during online training.
Reminders	If a system reminds users of their target behavior, the users will more likely achieve their goals.	<ul style="list-style-type: none"> <i>I think just a prompt can definitely make something - there's so many strategies and things that we're trying to remember. If someone else could remind you or prompt you, that's one less thing to try and remember.</i> [ID number 6, female, occupational therapist] <i>Think it could be a pain in the neck [Daily test message or alert].</i> [ID number 13, female, GP^b] 	<ul style="list-style-type: none"> A reminder to take a break to recover.
Suggestion	Systems offering fitting suggestions will have greater persuasive powers.	<ul style="list-style-type: none"> <i>I guess the app can also, sort of reminders around how to stop comparing yourself to others, write down goals, celebrate your small wins. Those sorts of things. Not necessarily about, "Oh you're doing a great job!" It's more about, "Don't forget to be mindful in what you're doing but be practical at the same time". " 'o, they're the "sor's of things that really appeal to me. Yeah.</i> [ID number 16, female, nurse educator] <i>If you could put a link to something that you could go - say for example, "You're more capable when you're relaxed. Here's a two-minute relaxation period", whatever it is. Use two minutes of a relaxation tape.</i> [ID number 16, female, nurse educator] 	<ul style="list-style-type: none"> Suggestions that the users are more capable when they are relaxed with a link to a 2-minute relaxation video.
Similarity	People are more readily persuaded through systems that remind them of themselves in some meaningful way.	<ul style="list-style-type: none"> <i>And I loved linking in to the rural webinars, conferences, that have happened. Because I'm listening to all the, I would have to say, fantastic innovations that some of these clinicians have come up with, and it is inspiration to you as a clinician. It's like, "Oh, I can actually do that", or, "I have the capacity to do this.</i> [ID number 17, female, occupational therapist] 	<ul style="list-style-type: none"> Using medical terminology for target audience. Health professionals lead discussion groups. Health professionals demonstrate innovations that others feel can be replicated.
Liking	A system that is visually attractive for its users is likely to be more persuasive.	<ul style="list-style-type: none"> <i>And I did the preventing dementia one, and that was fantastic because it was very professional, and it was multiple modules over four weeks. They did - very professionally presented videos, fancy PowerPoints, and they would have videos that would open up where they would have question and answer sessions with the researchers.</i> [ID number 6, female, occupational therapist] 	<ul style="list-style-type: none"> An app has professionally developed and visually attractive content.
Social role	If a system adopts a social role, users will more likely use it for persuasive purposes.	<ul style="list-style-type: none"> <i>We work on supporting each other that way. We just have conversations – we use Skype a lot rather than emails. Microsoft Teams I think is going to be the next step, but we actually really like that. It's a bit like being able to pop into a room and just go, "What do you think about this? or, 'Can you tell me what I could do about this?' without having to have all the formality of an email.</i> [ID number 9, female, physiotherapist] 	<ul style="list-style-type: none"> An app that has communities of practices or social events. An organization supports the members if in need.

^aSource: Oinas-Kukkonen and Harjumaa Marja [17].^bGP: general practitioner.

Table 3. Persuasive design techniques—definitions and implementation examples for developing a health workforce capability app. Domain: System Credibility Support.

System credibility support	Persuasive design technique definition ^a	Quote qualitative interviews health professionals	Implementation examples
Trustworthiness	A system that is viewed as trustworthy will have increased powers of persuasion.	<ul style="list-style-type: none"> • <i>As long as there's a privacy umbrella over it, that's acceptable.</i> [ID number 7, male, pain specialist] • <i>When I think about those things [privacy], I do feel concerned. But we're just so enmeshed in it that it's like it's too late. So I feel like what's the point of worrying about it, it's just another thing to worry about.</i> [ID number 15, male, GP^b] 	<ul style="list-style-type: none"> • App provides links to reputable websites such as Beyond Blue. • Privacy statements.
Expertise	A system that is viewed as incorporating expertise will have increased powers of persuasion.	<ul style="list-style-type: none"> • <i>They [the developers] need to have extensive clinical knowledge and they have to update it on a regular basis. So, Russel Harris is, of course, that's different. But, but the GPnotebook is regularly updated, so you feel you can rely on it. It uses guidelines. And they research and they make sure that they're following that.</i> [ID number 13, female, GP] • <i>And also to collect clinically-actionable data that's important data that we know makes a difference to the outcomes.</i> [ID number 7, male, pain specialist] 	<ul style="list-style-type: none"> • App provides extensive clinical knowledge based on latest evidence and specialist contributions.
Surface credibility	People make initial assessments of the system credibility based on a firsthand inspection.	<ul style="list-style-type: none"> • <i>It doesn't look so professional when your phone's going, "buzz, buzz, buzz" on your desk the whole time.</i> [ID number 15, male, GP] • <i>Like pharmaceutical ads...I think it's going to the dark side really a bit.</i> [ID number 13, female, GP] 	<ul style="list-style-type: none"> • App is updated regularly and there are no failing links or out-of-date information. • No or carefully selected commercial ads. • Content portion of the app is derived from reputable sources and relevant credentials of authors are displayed prominently.
Real-world feel	A system that highlights people or organization behind its content or services will have more credibility.	<ul style="list-style-type: none"> • <i>Absolutely, if I was feeling really low and I thought that someone would ring and check on me, absolutely, I think that would be a really good thing to do. I'd be okay with that.</i> [ID number 9, female, physiotherapist] 	<ul style="list-style-type: none"> • App provides information about the organization or rural health professionals or both. • App supports members to contact real people within the organization. • Organization provides real-time support to improve health workforce capability.
Authority	A system that leverages roles of authority will have enhanced powers of persuasion.	<ul style="list-style-type: none"> • <i>So I probably tap into all the pillar organisations that provide training as well.</i> [ID number 9, female, physiotherapist] • <i>Yes, and I wish that we have something like this which is Australian, with Australian guidelines or something like that, because mostly of course the guidelines will be American guidelines, but I will have a quick idea and then try to find the Australian guidelines to match those.</i> [ID number 2, male, GP] 	<ul style="list-style-type: none"> • Australian government website reference or guidelines. • Link to an official government-recognized network.
Third-party endorsements	Third-party endorsements, especially from well-known and respected sources, boost perceptions on system credibility.	<ul style="list-style-type: none"> • <i>I kind of am tapping more into < Government Network > that I work with as well. So I'm part of that and that's a really great resource, and there's another couple of rural physios and we keep a voice for rural health in that space and also support each other.</i> [ID number 9, female, physiotherapist] 	<ul style="list-style-type: none"> • An official government agency has endorsed the app and this is displayed on the app.
Verifiability	Credibility perceptions will be enhanced if a system makes it easy to verify the accuracy of site content via outside sources.	<ul style="list-style-type: none"> • <i>I think things that have a specific, kind of Australian – specific information that's definitely going to be relevant to what our local health services can provide and is available locally is important as well.</i> [ID number 15, male, GP] 	<ul style="list-style-type: none"> • Offer access to Australian-based information. • References lists and clear links to original sources.

^aSource: Oinas-Kukkonen and Harjumaa Marja [17].^bGP: general practitioner.

Table 4. Persuasive design techniques—definitions and implementation examples for developing a health workforce capability app. Domain: Social Support.

Social support	Persuasive design technique definition ^a	Quote qualitative interviews health professionals	Implementation examples
Social learning	A person will be more motivated to perform a target behavior if (s)he can use a system to observe others performing the behavior.	<ul style="list-style-type: none"> • <i>whatever we can get into in terms of progression or development and try and share that amongst ourselves and invite each other to that kind of gatherings as well.</i> [ID number 10, female speech pathologist] • <i>I mean I certainly when I'm travelling to other locations for work, then I use podcasts for educational purposes, which are extremely useful while I'm driving. Now on a regular basis <major urban centre> unit run for just 30 minutes a weekly professional development, where clinicians throughout the state present on a particular topic. So you've got the <small town> therapist presenting, you've got the <small town> therapist presenting, and everyone has an opportunity to throw their ideas, have group discussion. Fantastic, and that happens at the same time every week. For professional development it's really good, and that professional isolation, it's an excellent tool.</i> [ID number 17, female, occupational therapist] 	<ul style="list-style-type: none"> • Online meetings with rural health professionals to discuss health workforce capability. • Use of podcasts to listen to other health professionals when driving long distances. • Display number of views for content. • Allow feedback and comments on content. • A discipline-specific unit (eg, occupational therapy) coordinates weekly professional development sessions, occurring at the same time, which are accessible to all clinicians across the whole state. The coordination unit can be set up in a rural area rather than a city.
Social comparison	System users will have a greater motivation to perform the target behavior if they can compare their performance with that of others.	<ul style="list-style-type: none"> • <i>We do cross professional stuff too, so if there's physios that have come across something they might forward to you.</i> [ID number 10, female speech pathologist] • <i>"We just have conversations – we use Skype a lot rather than emails. Microsoft Teams I think is going to be the next step, but we actually really like that. It's a bit like being able to pop into a room and just go, "What do you think about this? or, 'Can you tell me what I could do about this?' without having to have all the formality of an email.</i> [ID number 13, female, GP^b] 	<ul style="list-style-type: none"> • Use chatrooms to allow for real-time discussions to compare how to improve health workforce capability. • Digital badges and milestones (eg, 50 articles read).
Normative influence	A system can leverage normative influence or peer pressure to increase the likelihood that a person will adopt a target behavior.	<ul style="list-style-type: none"> • <i>I feel that at the moment, I could do with a bit of help in motivation. It would be really good to discuss things with somebody, just to kind of...Almost like a careers advisor, really, in my State or somebody who would sit and say, "Well, these are the options." Or perhaps give me new ideas.</i> [ID number 13, female, GP] 	<ul style="list-style-type: none"> • App provides access to career tips, and career advisors or coaches who may have a normative influence by increasing the likelihood of the health professional being motivated to work on their capability.
Social facilitation	System users are more likely to perform target behavior if they discern via the system that others are performing the behavior along with them.	<ul style="list-style-type: none"> • <i>Our team communicates about training opportunities.</i> [ID number 9, female, physiotherapist] 	<ul style="list-style-type: none"> • Allow colleagues to share and discuss online training opportunities easily.
Cooperation	A system can motivate users to adopt a target attitude or behavior by leveraging human beings' natural drive to cooperate.	<ul style="list-style-type: none"> • <i>The benefits that I find telehealth really good for and which is actually starting to really come about is linking in to professional developments and interest groups.</i> [ID number 17, female, occupational therapist] 	<ul style="list-style-type: none"> • App is linking in to collaborative professional development activities and interest groups. • App allows creation of community of practice. • Community of practice permits creation of subgroups based on interest.

Social support	Persuasive design technique definition ^a	Quote qualitative interviews health professionals	Implementation examples
Competition	A system can motivate users to adopt a target attitude or behavior by leveraging human beings' natural drive to compete.	<ul style="list-style-type: none"> Like you can do the team things on that <exercises> and they're much more positive in, You've got to do more than – to beat your whatever in the team. [ID number 7, male, pain specialist] 	<ul style="list-style-type: none"> Key performance indicator tracker for individual tracking of exercise or work activities or other personalized set goals that improve health workforce capability over a period (competition with self or others). Digital badges. Recognition of best practices. Mini-competitions.
Recognition	By offering public recognition for an individual or group, a system can increase the likelihood that a person/group will adopt a target behavior.	<ul style="list-style-type: none"> So that recognition of skills and being valued is a huge one for senior therapists. [ID number 17, female, occupational therapist] 	<ul style="list-style-type: none"> Published stories of people being publicly recognized to demonstrate members are being valued for displaying capability. Digital badges.

^aSource: Oinas-Kukkonen and Harjumaa Marja [17].

^bGP: general practitioner.

Table 5. Persuasive design techniques—definitions and implementation examples for developing a health workforce capability app. Domain: Cialdini's Principles.

Cialdini	Persuasive design technique definition ^a	Quote qualitative interviews health professionals	Implementation examples
Commitment/consistency	Are a pair of interrelated attributes in the sense that people often adhere (consistently) to their significant choices (commitments).	And if there was an app that actually tracked what you were doing, tracked your KPIs, then that information could be funnelled back through middle management, but also the higher levels as well [ID number 17, female, occupational therapist]	<ul style="list-style-type: none"> App allows for registering own KPIs^b (commitment) with set periods (eg, weekly, monthly [consistency]). KPIs can be communicated to supervisors and higher as requested.
Scarcity	Causes people to almost panic out of the fear that something will disappear or become unavailable, so they make an intent effort to acquire or preserve it.	— ^c	<ul style="list-style-type: none"> Circulation of grant opportunities with an emphasis on deadlines.
Social proof	Explains the human tendency to look around at others in society for reinforcement and direction in taking action.	But I was thinking...that when we couldn't save the little boy that died in front of our eyes, could digitally, or through Zoom or something, could we have gotten a trauma counsellor, instead of us driving over to <rural town> , a 200 Ks. And I didn't want to drive that day, but I needed to debrief. [ID number 3, female, remote nurse]	<ul style="list-style-type: none"> Shows number of members in a chat group or in a specific geographical location. Online support to assist in emergencies.
Reciprocity	Describes a human desire to make others feel appreciated by responding in ways that return good gestures.	So it's also very interesting just clinically – seeing what people are doing or their approaches to similar cases, so we encourage that kind of discussion a lot. [ID number 10, female speech pathologist]	<ul style="list-style-type: none"> People in a community of practice help each other with clinical problems or capability-related issues.

^aSource: Oyebode et al [20].

^bKPI: key performance indicator.

^cNo matching quote or data found.

Objectives

This study aims to:

1. explore rural health professionals' perceptions of health workforce capability and persuasive design techniques;

2. translate these into recommendations for designing a health workforce capability app to increase their impact and usefulness.

Methods

Data Collection

A qualitative analysis of interview data used in the 2 previous studies was conducted to evaluate how persuasive design techniques can be used to build a health workforce capability app. As outlined in Ramsden et al [10], 13 rural health professionals were interviewed about their understanding of health workforce capability, their perceived needs to improve health workforce capability, and how technological solutions can assist in improving rural health workforce capability and intentions to remain in rural practice.

Interview questions were informed by the work of Anderson and co-workers [22] and Jeffrey and colleagues [23] in terms of focusing on technology acceptance and behavior change. Persuasive strategy questions were informed by the work of Oyeboode and co-workers [20] as described above.

Recruitment occurred via the telephone, in person, or through the Rural Health Pro newsletter. Rural Health Pro is a digital platform that links health professionals and organizations interested in rural health [24]. A plain language statement was emailed to participants expressing interest. An interview time was established after consent. Telephone interviews were conducted by RR and SP, digitally recorded, and transcribed. Transcripts were not returned to participants for comment, verification, or correction. Upon transcription, identifying information was removed. The transcribed interviews were managed in MS Word. Both RR and SP have extensive experience in conducting qualitative interviews. SP has lived rurally since 2005 and worked in rural health research since 2006. RR has worked in rural health since 2012. These experiences have shaped the interviewers' desire to improve rural health workforce.

Participants included general practitioners (GPs), a pain specialist, nurses, and allied health practitioners. The group comprised 4 males and 9 females. Age ranged between 39 and 65 years, with an average age of 51 years. Interviews took on average 46 minutes (range 29-98 minutes). The interviewers had no previous existing relationship with the participants. RR and SP undertook reflexivity exercises during the data collection to ensure rigor, in addition to the checks during analysis listed below. RR and SP recognized that they have an interest in improving rural health workforce capability using technology. Both ethical and practical issues that arose during the interviews were discussed between the 2 interviewers to ensure alignment and rigor when conducting the interviews.

Analyses

Our systematic review of persuasive apps that are related to health workforce capability generated examples of how the various techniques can be used in practice to develop a health workforce capability app. Subsequently, the 32 persuasive techniques [11,17-19] were used as a framework to deductively analyze the data. Descriptions of the 32 persuasive design techniques are listed in Tables 1-5.

Verbatim transcripts were coded manually. The first 2 recordings were coded separately by 2 authors (KP and RR). The remaining

11 interviews were coded primarily by authors KP and SP. RR and AT were involved in coding transcripts and reaching consensus. SP and KP subsequently used the coded data to analyze and identify quotes to support the development of suggestions for how persuasive design techniques can be used to build a health workforce capability digital tool. The findings were discussed and checked by RR and AT to ensure validity of the data. JB provided specific feedback on user design and discussions were held to further shape the data interpretation.

Ethical Approval

The study was approved by the North Coast NNSWLHD Human Research Ethics Committee (2020/ETH03020).

Results

Tables 1-5 show how health professionals' needs and suggestions are reflected in the 32 persuasive design techniques that can be used when building a health workforce capability online support tool.

The results are displayed by the main domains in persuasive design techniques: Primary Task Support, Dialog Support, System Credibility Support, Social Support, and Cialdini's Principles of Persuasion. The researchers analyzed the qualitative data to identify how health professionals' perspectives could be translated into recommendations for designing a health workforce capability app to increase their persuasiveness. The majority of the 32 persuasive design techniques were reflected in the quotes from the interviews and led the researchers to a recommendation as to how these could be translated into practice, with the exception of competition and scarcity.

Primary Task Support techniques, for example, may be utilized to guide health professionals through a clinical protocol or professional development activity by reducing complex behaviors into simple tasks, tailoring evidence-based information to needs, and personalizing content (eg, Australian guidelines). Further, techniques such as simulation would enable health professionals to observe other clinicians performing procedures and rehearse behaviors themselves, accommodating health professionals' ideas of how they would like to use an app to build capability.

Health professional perspectives also highlighted the importance of credibility for app design. For example, under the domain System Credibility Support (Table 3) a GP mentioned "Like pharmaceutical ads...I think it's going to the dark side really a bit" [ID number 13, female, GP], suggesting that consideration should be given to developing criteria about which type of ads would be acceptable. Simultaneously, given the importance placed by the participants on addressing various social needs to improve health workforce capability, designers would need to consider eligibility criteria for joining online communities to safeguard the quality of online community members (System Credibility Support).

Giving feedback to support health professionals to move toward their goals, such as rewards and praise in the form of continuing professional development points or encouragement for

goal-directed behaviors, is a suggested design feature. Further, reminders about target behavior, such as taking a break, and suggestions related to capability-supporting behaviors that are suitable or appropriate to health professionals were examples provided that align with Dialog Support techniques.

Further analyses revealed that health professional needs and digital solutions mapped against persuasive techniques broadly align with the factors affecting health workforce capability. The analyses are summarized in Table 6. Based on health practitioners' insights and suggestions, practical recommendations have also been presented in Table 6. It is noted that the findings are not necessarily mutually exclusive

given the complexity. For example, depending on the level and governance, the recommendation about communities of practices on team level can also be viewed as Primary Task Support and Social Support if, for example, a national or state body would manage the communities of practice. The recommendation relating to organizational level refers to the fact that referrals and closed-loop communication are professional activities that form part of the practitioner's primary scope of work. Thus, this function would be simplifying a process that already takes place and falls within the definition of reduction. The design techniques appear acceptable to rural health professionals and can be incorporated into future apps that focus on improving or maintaining health workforce capability across several areas.

Table 6. Factors influencing health workforce capability mapped against persuasive design techniques (design feature category).

Factors influencing health workforce capability and persuasive design technique	Persuasive design feature domain	Recommendations provided by participants that can be linked with health workforce capability factors
Health and personal qualities		
<ul style="list-style-type: none"> Self-monitoring 	Primary Task Support	<ul style="list-style-type: none"> A reminder on an app to take a break to recover.
<ul style="list-style-type: none"> Reminder 	Dialog Support	<ul style="list-style-type: none"> Improved fitness through an exercise monitoring and scheduling app.
Competencies and skills		
<ul style="list-style-type: none"> Expertise 	System Credibility Support	<ul style="list-style-type: none"> An app that provides extensive clinical knowledge based on latest evidence and specialist contributions. Clinical competence building through online education.
Values, attitudes, and motivation		
<ul style="list-style-type: none"> Surface credibility 	System Credibility Support	<ul style="list-style-type: none"> Increased credibility by not displaying pharmaceutical advertisements on apps.
<ul style="list-style-type: none"> Recognition Social learning Social facilitation 	Social Support	<ul style="list-style-type: none"> Published stories of people being publicly recognized for work they have done to demonstrate members are being valued.
Factors outside of work		
<ul style="list-style-type: none"> Social comparison Social learning 	Social Support	<ul style="list-style-type: none"> Using chatrooms to allow for real-time discussions and to see other rural health professionals taking a holiday.
Team level		
<ul style="list-style-type: none"> Reciprocity 	Cialdini's principles	<ul style="list-style-type: none"> Health professionals participating in a community of practice and helping each other with clinical problems or capability-related issues.
Organizational level		
<ul style="list-style-type: none"> Reduction 	Primary Task Support	<ul style="list-style-type: none"> GPs^a being able to book patient appointments for allied health professionals online and that which link back to the GP so both parties receive reports. This could potentially work both ways. Allied health professionals being able to book appointments with the GP for their patients.
Labor market		
<ul style="list-style-type: none"> Tailoring 	Primary Task Support	<ul style="list-style-type: none"> Job vacancies for rural health professionals are tailored to health care professionals' interest and discipline.

^aGP: general practitioner.

Discussion

Principal Findings

A qualitative analysis of semistructured interviews with rural health professionals was undertaken to investigate the alignment of user perceptions of health workforce capability with persuasive design techniques. These findings were then translated into recommendations for designing a health workforce capability app to increase their impact and usefulness (Tables 1-6). The authors found that the persuasive design techniques can be applied across the factors that influence health workforce capability including health and personal qualities, competencies and skills, values, attitudes, and motivation, as well as factors that operate outside of work, and at the team, organizational, and labor market levels. There is alignment between the needs of health professionals and persuasive design techniques. The health professionals' interviews clearly identified the persuasive design techniques that were appealing. Many suggestions link back to the need for credible, tailored, and localized solutions such as having Australian-based evidence, localized communities of practice, access to learning from other rural health professionals, and referral pathways to other clinicians.

It was also clear that rural health professionals in our study aspired to have more efficient systems to improve patient access to care, quality care, and to enable working in interprofessional team-based care. Persuasive design techniques that can assist here are, for example, tailoring, reduction, social roles of digital technology, and authority. These techniques can influence team and organizational factors to improve health workforce capability. Indeed, the need for team-based care is an old adage, and barriers and facilitators to this have been explored over time [25]. Notably, our previous work [10] demonstrated that COVID-19 has accelerated the capability and willingness for team-based care arrangements in primary and remote care through using digital solutions. This study has clear recommendations as to how persuasive design techniques can be used in digital solutions to further improve patient access, quality care, and promote team-based care. For example, one GP recommended that GPs should be able to book an appointment with a local allied health professional with an online booking system (Primary Task Support) so the patient does not have to do that. This ensures the patient is booked in (tunneling) and reduces the risk of the patient not following up with their allied health appointment, thereby enhancing the quality of care as well as the GPs own health workforce capability. Ways in which this system may improve a GPs and allied health professional's workforce capability are reduced red tape and less inconvenience in making referrals; improved close loop communication that reduces the workload for the GP and the allied health professional; better outcomes for patient that leads to further reduction in the future workload; positive feedback to GPs and allied health professionals (recognition); and improved job satisfaction for both allied health professionals and GPs as closer relationships may be developed overtime.

Overall, health professionals were open to self-monitoring of their work performance and some already regularly use

wearables to monitor their health. Although using wearables to measure work performance and stress levels at work has potential [26], it is still in its infancy. A 2019 randomized controlled trial showed that using wearables to improve emergency physicians' well-being through monitoring their pulse rate while at work was feasible but not recommended [27]. The authors found that the biosensor could not provide reliable estimates of metrics of interest in their study context. However, Ferdous and colleagues [28] found that patterns of smartphone app usage were correlated with stress levels in work environments and thus recommended that these could be used to measure stress at work. Alhasani et al [29] presented the possibility of an app that creates behavioral data, based on self-report and sensors, that can be analyzed in real-time to predict users' needs and provide tailored interventions for the user. In the case of rural health workforce capability, if a health professional's heart rate variability indicates high stress levels derived from a sensor, the app can recommend a predetermined meditation session. The behavioral data can also be merged with self-reports generated by in-app journals or self-report. Synthesizing and interpreting this information using machine learning could improve the accuracy of predictions and hence provide more targeted and personalized recommendations for the user.

Social inclusiveness is a known contributing factor for improved health workforce capability [10]. Therefore, computer-mediated persuasion [17] using digital social communication tools that have a real-world feel and demonstrate that there are real people and experts (surface credibility) behind a digital tool or content is important. Health professionals will be more likely to engage with a digital system if they know that real people and respected experts are involved. In a study of nurses, Mayer and colleagues [30] found that approximately half of the nurses used apps for their work and stressed the importance of apps being validated by credible bodies before they can be used in practice. Expert design, content, and involvement will lead to greater system credibility and also improved social connections and support channels for rural health professionals. This stresses the importance of policies to ensure high-quality membership of digital communities and, thus, system credibility (persuasive design technique) to guarantee impact and engagement with the tool. Ultimately, the quality of the online membership can have an impact on usage and engagement of health workforce capability apps. Therefore, factors to consider when developing an app are closed or open membership; verification of credentials; potential publication of credentials; role and credentials of moderators; and rules of participation.

Validation of Persuasive Systems Design Model and Persuasive Design Theories

Our work has contributed to the further validation of the Persuasive Systems Design model [17] through empirically aligning elements of the model to increase persuasive system content and functionality with real-world applied data, in this case the health workforce capability of rural health professionals. We analyzed the *use context* through identifying problem domain-dependent features which are the factors that influence health workforce capability [10]. We also analyzed the *user context* that are the user-dependent features such as goals or

motivations of health professionals [10]. Subsequently, we analyzed the *technology context* through identifying the technology-dependent features for health workforce capability-related apps [11]. In this paper, we then described example software apps and implementations. Some of these examples are relatively novel in addressing health workforce capability needs such as geolocation of health professionals in a rural area to reduce social isolation [10] and some are already used in varying capability building-related apps such as easy access to clinical guidelines to increase medical performance [31] in apps such as UptoDate [11] or online communities such as Rural Health Pro [24]. Rural Health Pro is a digital platform that connects people and organizations who care about keeping rural communities healthy. Notably, we found that sometimes comments made by health professionals could apply to several persuasive design techniques. This is confirmed by feedback from user experience designers that, for example, tailoring and personalization are often the same in practice.

Study Strengths and Limitations

Overall, the alignment between the needs of health professionals and persuasive design techniques to improve health workforce capability through optimizing the user experience means that we can support rural health professionals which may lead to greater retention [9,11,13]. The above findings are important as they confirm that these techniques can be used to develop a future prototype that may assist health professionals in improving or maintaining their health workforce capability which in turn may increase recruitment and retention in rural areas. However, we stress that it is unlikely that a single app could cover all persuasive design techniques explored in this paper, nor do we imply a health workforce capability app should do so. For example, if we built a health workforce capability app for GPs, it likely would not have the ability to book appointments for allied health professionals as that may not make sense from a product or commercial perspective to bundle and market those features together. Additionally, a system itself needs to provide enough value and solve a problem for health professionals, regardless of the number of persuasive techniques built into the app. A very useful system without any or a limited number of persuasive techniques could still improve health professional capability. Persuasive design techniques are only one element to consider in the development process of a health workforce capability app. For example, usability heuristics are useful to develop an easy-to-use interface, marketing campaigns may be necessary to promote uptake of a health workforce capability app, and service blueprinting can be used to guide the whole process.

Ethical Considerations

It is important to consider the ethical factors in using persuasive strategies. Jacobs [32] explored the use and ethical concerns of persuasive technology for vulnerable people. While health professionals are not considered vulnerable overall, they can be vulnerable when working in isolation for long periods and this may be exacerbated by the COVID-19 pandemic and other natural disasters such as bushfires, droughts, and floods. Their strong sense of responsibility and commitment to their patients and community make health professionals potentially more

vulnerable over time as their energy levels and resources are depleted [33]. However, the point of using persuasive design techniques in digital solutions is to make health professionals feel more capable when they are feeling overstretched or underconfident. The implication is that thorough consideration of, and trialing of, appropriate strategies is important.

Vulnerability can be caused by intrinsic factors (within a person) or can be caused by external factors that are situational specific [32]. An example of situational vulnerability is that an app may be designed for support but reduces the health professional's privacy, and increases feelings of being watched. It may also increase their feelings of anxiety about the potential for negative consequences if their capability goes down, and they could potentially start to feel powerless. The ethical concept of autonomy may be violated in this example.

The design should ensure that the user, in this case, the health professional, can easily retract their consent to take part to avoid coercion [32]. This should include the ability to withdraw any data already submitted or collected about them. This will have the secondary function of improving engagement as the participants are assured of the ability to cease their involvement at any time. For example, a rural health professional when highly capable may be willing to partake in an app that measures their capability; however, if they feel less capable for any reason, it may well be that they become increasingly vulnerable. Completing a rating on a regular basis can make people aware of their shortcomings and have unintended consequences [32], such as exacerbated feelings of incompetence or reinforcing unconscious incompetence.

Jacobs [32] recommends that to safeguard against unintended consequences, technology designers should understand the experiences, interests, and needs of prospective users through inclusion of vulnerable populations at every stage of the design process. Our study described here has done that and is the first step in the co-design process. Designers need to ensure they understand the values, needs, and interests that are important to the users. A second safeguard is to take into account real-life contexts, which can reduce unintended negative consequences [32]. The latter issue demonstrates the importance of considering a digital tool as part of an integrated approach that includes a real-life support component. Finally, there must be a clear and continuous feedback mechanism where users can identify and raise issues of concern for the developers to remedy.

Organizations developing or managing apps or online communities would need to ensure high-quality online membership. Ethical principles, membership criteria, and policies will need to be developed to ensure mechanisms are in place to ensure members or users genuinely represent the purpose and values of the community.

To design a digital health workforce capability solution, it is important to select and apply an ethical framework that fits the health workforce capability needs and assists in carefully weighing up risks and benefits. Many ethical frameworks exist in the medical research field. They can be disease specific or goal specific such as an ethical framework for COVID-19 contact tracing [34] or more general such as the work conducted by Tokgöz and co-workers [35] who developed an ethical

framework for health and medical apps based on a systematic review and expert interviews. However, it is ultimately recommended to use an ethical framework that aligns with international standards. Several international standards are available and selection of a standard will depend on the context of the app such as 'ISO/TS 82304-2:2021 Health software — Part 2: Health and wellness apps — Quality and reliability' [36] or 'the ISO/TS 17033:2019 Ethical claims and supporting information — Principles and requirements' [37]. The latter can be used when specific standards are not available or can complement existing standards. It covers principles and requirements for developing and declaring ethical claims.

Given that rural health professionals are often subject to burn out due to isolation, high stress, long working hours, lack of local staff, and limited resources, being able to provide a technology solution with persuasive strategies to boost capability is a positive aspect.

Conclusions

Rural health professionals' perceptions of health workforce capability informed which persuasive design techniques can be used to optimize the user experience of an app. These were translated into recommendations for designing a health workforce capability app to increase likelihood of adoption. This study has also contributed to the further validation of the Persuasive Systems Design model through empirically aligning elements of the model to increase persuasive system content and functionality with real-world applied data, in this case the health workforce capability of rural health professionals. Our findings confirm that these techniques can be used to develop a future prototype of an app that may assist health professionals in improving or maintaining their health workforce capability, which in turn may increase recruitment and retention in rural areas. Health professionals need to be central during the design phase. Interventions are needed to provide a supportive environment to rural and remote health professionals to increase their rural health workforce capability.

Acknowledgments

We thank the NIB Foundation for contributing part funding for the analyses of this study. The funder has had no impact on the study design, interpretation, or outcomes of this study. We thank Katie Carlisle for contributing to study design and project management.

Authors' Contributions

SWP, RR, and RC designed the study. SWP, RR, and KP analyzed the qualitative data, with input from AJHT. SWP, RR, AJHT, and KP interpreted study results with input from RC, JB, BE, and ME. AJHT drafted the introduction. SWP drafted the remainder of the initial manuscript with major inputs from RR, RC, AJHT, and KP. JB, ME, and BE provided further content expertise. All authors contributed to developing the study materials, writing the manuscript, and read and approved the final manuscript.

Conflicts of Interest

New South Wales Rural Doctors Network and authors declare an interest in developing the health workforce capability of rural health professionals.

References

1. Angrish S, Sharma M, Bashar M, Tripathi S, Hossain M, Bhattacharya S, et al. How effective is the virtual primary healthcare centers? An experience from rural India. *J Family Med Prim Care* 2020 Feb;9(2):465-469 [FREE Full text] [doi: [10.4103/jfmpc.jfmpc_1124_19](https://doi.org/10.4103/jfmpc.jfmpc_1124_19)] [Medline: [32318365](https://pubmed.ncbi.nlm.nih.gov/32318365/)]
2. Streeter RA, Snyder JE, Kepley H, Stahl AL, Li T, Washko MM. The geographic alignment of primary care Health Professional Shortage Areas with markers for social determinants of health. *PLoS One* 2020;15(4):e0231443 [FREE Full text] [doi: [10.1371/journal.pone.0231443](https://doi.org/10.1371/journal.pone.0231443)] [Medline: [32330143](https://pubmed.ncbi.nlm.nih.gov/32330143/)]
3. Walsh S, Lyle D, Thompson S, Versace V, Browne L, Knight S, et al. The role of national policies to address rural allied health, nursing and dentistry workforce maldistribution. *Medical Journal of Australia* 2020;213(11):S18-S22 [FREE Full text]
4. Fitzpatrick M, Garsia K, Eyre K, Blackhall C, Pit S. Emotional exhaustion among regional doctors in training and the application of international guidelines on sustainable employability management for organisations. *Aust Health Rev* 2020 Aug;44(4):609-617. [doi: [10.1071/AH19121](https://doi.org/10.1071/AH19121)] [Medline: [32183938](https://pubmed.ncbi.nlm.nih.gov/32183938/)]
5. Xu W, Pan Z, Li Z, Lu S, Zhang L. Job Burnout Among Primary Healthcare Workers in Rural China: A Multilevel Analysis. *Int J Environ Res Public Health* 2020 Jan 22;17(3):727 [FREE Full text] [doi: [10.3390/ijerph17030727](https://doi.org/10.3390/ijerph17030727)] [Medline: [31979154](https://pubmed.ncbi.nlm.nih.gov/31979154/)]
6. NSW Rural Doctors Network. 2020-21 Primary health workforce needs assessment. Summary Report. NSW Rural Doctors Network. 2021. URL: https://www.nswrdn.com.au/client_images/2268722.pdf [accessed 2022-03-21]
7. Buyx P, Humphreys J, Wakerman J, Pashen D. Systematic review of effective retention incentives for health workers in rural and remote areas: towards evidence-based policy. *Aust J Rural Health* 2010 Jun;18(3):102-109. [doi: [10.1111/j.1440-1584.2010.01139.x](https://doi.org/10.1111/j.1440-1584.2010.01139.x)] [Medline: [20579020](https://pubmed.ncbi.nlm.nih.gov/20579020/)]

8. Yang BK, Johantgen ME, Trinkoff AM, Idzik SR, Wince J, Tomlinson C. State Nurse Practitioner Practice Regulations and U.S. Health Care Delivery Outcomes: A Systematic Review. *Med Care Res Rev* 2021 Jun;78(3):183-196. [doi: [10.1177/1077558719901216](https://doi.org/10.1177/1077558719901216)] [Medline: [31997710](https://pubmed.ncbi.nlm.nih.gov/31997710/)]
9. Martiniuk AL, Colbran R, Ramsden R, Edwards M, Barrett E, O'Callaghan E, et al. Capability ... what's in a word? Rural Doctors Network of New South Wales Australia is shifting to focus on the capability of rural health professionals. *Rural Remote Health* 2020 Jul;20(3):5633 [FREE Full text] [doi: [10.22605/RRH5633](https://doi.org/10.22605/RRH5633)] [Medline: [32650644](https://pubmed.ncbi.nlm.nih.gov/32650644/)]
10. Ramsden R, Pit S, Payne K, Tan A, Eames B, Edwards M, et al. Development of a framework to promote rural health workforce capability through digital solutions: A qualitative study of user perspectives. *Digital Health* 2022;1 (forthcoming)(forthcoming).
11. Pit SW, Tan AJH, Ramsden R, Payne K, Freihaut W, Hayes O, et al. Persuasive Design Solutions for a Sustainable Workforce: Review of Persuasive Apps for Real-Time Capability Support for Rural Health Care Professionals. *JMIR Mhealth Uhealth* 2022 Feb 7;10(2):e33413 [FREE Full text] [doi: [10.2196/33413](https://doi.org/10.2196/33413)] [Medline: [35129447](https://pubmed.ncbi.nlm.nih.gov/35129447/)]
12. Fisk M, Livingstone A, Pit SW. Telehealth in the Context of COVID-19: Changing Perspectives in Australia, the United Kingdom, and the United States. *J Med Internet Res* 2020 Jun 09;22(6):e19264 [FREE Full text] [doi: [10.2196/19264](https://doi.org/10.2196/19264)] [Medline: [32463377](https://pubmed.ncbi.nlm.nih.gov/32463377/)]
13. Ramsden R, Colbran R, Christopher E, Edwards M. The role of digital technology in providing education, training, continuing professional development and support to the rural health workforce. *Health Education* 2021 Feb 18;122(2):126-149 [FREE Full text] [doi: [10.1108/he-11-2020-0109](https://doi.org/10.1108/he-11-2020-0109)]
14. Stoyanov SR, Hides L, Kavanagh DJ, Zelenko O, Tjondronegoro D, Mani M. Mobile app rating scale: a new tool for assessing the quality of health mobile apps. *JMIR Mhealth Uhealth* 2015;3(1):e27 [FREE Full text] [doi: [10.2196/mhealth.3422](https://doi.org/10.2196/mhealth.3422)] [Medline: [25760773](https://pubmed.ncbi.nlm.nih.gov/25760773/)]
15. Holden RJ, Karsh B. The technology acceptance model: its past and its future in health care. *J Biomed Inform* 2010 Feb;43(1):159-172 [FREE Full text] [doi: [10.1016/j.jbi.2009.07.002](https://doi.org/10.1016/j.jbi.2009.07.002)] [Medline: [19615467](https://pubmed.ncbi.nlm.nih.gov/19615467/)]
16. Kim J, Park H. Development of a health information technology acceptance model using consumers' health behavior intention. *J Med Internet Res* 2012;14(5):e133 [FREE Full text] [doi: [10.2196/jmir.2143](https://doi.org/10.2196/jmir.2143)] [Medline: [23026508](https://pubmed.ncbi.nlm.nih.gov/23026508/)]
17. Oinas-Kukkonen H, Harjumaa M. Persuasive Systems Design: Key Issues, Process Model, and System Features. *Communications of the Association for Information Systems* 2009;24(1):28. [doi: [10.17705/1CAIS.02428](https://doi.org/10.17705/1CAIS.02428)]
18. Cialdini RB. Harnessing the science of persuasion. *Harvard Business Review* 2001;79(9):72-81 [FREE Full text]
19. Hines J. An Analysis of the Use of Robert B. Cialdini's Six Persuasive Techniques in YouTube Videos about Endurance Training/Life Balance. Doctoral Dissertation. Edwardsville, IL: Southern Illinois University; 2021. URL: https://media.proquest.com/media/hms/PFT/2/CPzXJ?_s=gqIPU5iZCw7ZnZPt0YTvgJUKdg0%3D [accessed 2022-03-21]
20. Oyebo O, Ndulue C, Alhasani M, Orji R. Persuasive Mobile Apps for Health and Wellness: A Comparative Systematic Review. In: Gram-Hansen S, Jonassen T, Midden C, editors. *Persuasive Technology. Designing for Future Change. PERSUASIVE 2020. Lecture Notes in Computer Science*, vol 12064. Cham, Switzerland: Springer; 2020:163-181.
21. Carolan S, Harris PR, Cavanagh K. Improving Employee Well-Being and Effectiveness: Systematic Review and Meta-Analysis of Web-Based Psychological Interventions Delivered in the Workplace. *J Med Internet Res* 2017 Jul 26;19(7):e271 [FREE Full text] [doi: [10.2196/jmir.7583](https://doi.org/10.2196/jmir.7583)] [Medline: [28747293](https://pubmed.ncbi.nlm.nih.gov/28747293/)]
22. Anderson K, Burford O, Emmerton L. Mobile Health Apps to Facilitate Self-Care: A Qualitative Study of User Experiences. *PLoS One* 2016;11(5):e0156164 [FREE Full text] [doi: [10.1371/journal.pone.0156164](https://doi.org/10.1371/journal.pone.0156164)] [Medline: [27214203](https://pubmed.ncbi.nlm.nih.gov/27214203/)]
23. Jeffrey B, Bagala M, Creighton A, Leavey T, Nicholls S, Wood C, et al. Mobile phone applications and their use in the self-management of Type 2 Diabetes Mellitus: a qualitative study among app users and non-app users. *Diabetol Metab Syndr* 2019;11:84 [FREE Full text] [doi: [10.1186/s13098-019-0480-4](https://doi.org/10.1186/s13098-019-0480-4)] [Medline: [31636719](https://pubmed.ncbi.nlm.nih.gov/31636719/)]
24. NSW Rural Doctors Network. Rural Health Pro. NSW Rural Doctors Network. 2022. URL: <https://www.ruralhealthpro.org/> [accessed 2022-03-21]
25. Crompton D, Hsu C, Coleman K, Fishman PA, Liss DT, Ehrlich K, et al. Barriers and facilitators to team-based care in the context of primary care transformation. *J Ambul Care Manage* 2015;38(2):125-133. [doi: [10.1097/JAC.000000000000056](https://doi.org/10.1097/JAC.000000000000056)] [Medline: [25748261](https://pubmed.ncbi.nlm.nih.gov/25748261/)]
26. Hopkins L, Stacey B, Robinson DBT, James OP, Brown C, Egan RJ, et al. Consumer-grade biosensor validation for examining stress in healthcare professionals. *Physiol Rep* 2020 Jun;8(11):e14454 [FREE Full text] [doi: [10.14814/phy2.14454](https://doi.org/10.14814/phy2.14454)] [Medline: [32489016](https://pubmed.ncbi.nlm.nih.gov/32489016/)]
27. Peters GA, Wong ML, Joseph JW, Sanchez LD. Pulse Rate Variability in Emergency Physicians During Shifts: Pilot Cross-Sectional Study. *JMIR Mhealth Uhealth* 2019 Oct 2;7(10):e13909. [doi: [10.2196/13909](https://doi.org/10.2196/13909)] [Medline: [31579017](https://pubmed.ncbi.nlm.nih.gov/31579017/)]
28. Ferdous R, Osmani V, Mayora O. Smartphone apps usage patterns as a predictor of perceived stress levels at workplace. 2015 Presented at: 9th International Conference on Pervasive Computing Technologies for Health care (PervasiveHealth); May 20-23, 2015; Istanbul, Turkey p. 225-228. [doi: [10.4108/icst.pervasivehealth.2015.260192](https://doi.org/10.4108/icst.pervasivehealth.2015.260192)]
29. Alhasani M, Mulchandani D, Oyebo O, Orji R. A Systematic Review of Persuasive Strategies in Stress Management Apps. 2020. URL: http://ceur-ws.org/Vol-2662/BCSS2020_paper4.pdf [accessed 2021-09-29]
30. Mayer MA, Rodríguez Blanco O, Torrejon A. Use of Health Apps by Nurses for Professional Purposes: Web-Based Survey Study. *JMIR Mhealth Uhealth* 2019 Nov 01;7(11):e15195 [FREE Full text] [doi: [10.2196/15195](https://doi.org/10.2196/15195)] [Medline: [31682587](https://pubmed.ncbi.nlm.nih.gov/31682587/)]

31. Dijkstra R, Wensing M, Thomas R, Akkermans R, Braspenning J, Grimshaw J, et al. The relationship between organisational characteristics and the effects of clinical guidelines on medical performance in hospitals, a meta-analysis. *BMC Health Serv Res* 2006 Apr 28;6:53 [FREE Full text] [doi: [10.1186/1472-6963-6-53](https://doi.org/10.1186/1472-6963-6-53)] [Medline: [16646968](https://pubmed.ncbi.nlm.nih.gov/16646968/)]
32. Jacobs N. Two ethical concerns about the use of persuasive technology for vulnerable people. *Bioethics* 2020 Jun;34(5):519-526 [FREE Full text] [doi: [10.1111/bioe.12683](https://doi.org/10.1111/bioe.12683)] [Medline: [31617216](https://pubmed.ncbi.nlm.nih.gov/31617216/)]
33. Smyth J, Pit SW, Hansen V. Can the work ability model provide a useful explanatory framework to understand sustainable employability amongst general practitioners: a qualitative study. *Hum Resour Health* 2018 Jul 24;16(1):32 [FREE Full text] [doi: [10.1186/s12960-018-0292-x](https://doi.org/10.1186/s12960-018-0292-x)] [Medline: [30041658](https://pubmed.ncbi.nlm.nih.gov/30041658/)]
34. Lo B, Sim I. Ethical Framework for Assessing Manual and Digital Contact Tracing for COVID-19. *Ann Intern Med* 2021 Mar;174(3):395-400. [doi: [10.7326/m20-5834](https://doi.org/10.7326/m20-5834)] [Medline: [33076694](https://pubmed.ncbi.nlm.nih.gov/33076694/)]
35. Tokgöz P, Eger H, Funke S, Gutjar A, Nguyen-Thi T, Dockweiler C. Development and presentation of an ethical framework for health and medical apps. *Journal of the International Society for Telemedicine and eHealth* 2019;16(7):e15-e11 [FREE Full text]
36. International Organization for Standardization (ISO). ISO/TS 82304-2:2021. Health software — Part 2: Health and wellness apps — Quality and reliability. ISO. 2021. URL: <https://www.iso.org/standard/78182.html> [accessed 2022-03-21]
37. International Organization for Standardization (ISO). ISO/TS 17033:2019 Ethical claims and supporting information — Principles and requirements. ISO. 2021. URL: <https://www.iso.org/standard/29356.html> [accessed 2022-03-21]

Abbreviations

GP: general practitioner

HITAM: Health Information Technology Acceptance Model

KPI: key performance indicator

MARS: Mobile Application Rating Scale

PSD: Persuasive System Design Model

TAM: Technology Acceptance Model

Edited by G Eysenbach; submitted 20.11.21; peer-reviewed by I Wilson; comments to author 10.12.21; revised version received 11.12.21; accepted 28.12.21; published 02.05.22.

Please cite as:

Pit S, Ramsden R, Tan AJH, Payne K, Barr J, Eames B, Edwards M, Colbran R

Persuasive Design Techniques and App Design Recommendations to Improve Health Workforce Capability in Rural Health Professionals: What Do Users Want and How Does an App Help?

JMIR Hum Factors 2022;9(2):e35094

URL: <https://humanfactors.jmir.org/2022/2/e35094>

doi: [10.2196/35094](https://doi.org/10.2196/35094)

PMID: [35499866](https://pubmed.ncbi.nlm.nih.gov/35499866/)

©Sabrina Pit, Robyn Ramsden, Aaron JH Tan, Kristy Payne, James Barr, Benjamin Eames, Mike Edwards, Richard Colbran. Originally published in *JMIR Human Factors* (<https://humanfactors.jmir.org>), 02.05.2022. This is an open-access article distributed under the terms of the Creative Commons Attribution License (<https://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in *JMIR Human Factors*, is properly cited. The complete bibliographic information, a link to the original publication on <https://humanfactors.jmir.org>, as well as this copyright and license information must be included.

Original Paper

Exploring Human-Data Interaction in Clinical Decision-making Using Scenarios: Co-design Study

Helena Tendedez¹, BSc, PhD; Maria-Angela Ferrario², BSc, MSc, PhD; Roisin McNaney³, BSc, PhD; Adrian Gradinar⁴, BSc, PhD

¹School of Computing and Communications, Lancaster University, Lancaster, United Kingdom

²School of Electronics, Electrical Engineering and Computer Science, Queen's University Belfast, Belfast, United Kingdom

³Department of Human Centred Computing, Monash University, Melbourne, Australia

⁴Lancaster Institute for the Contemporary Arts, Lancaster University, Lancaster, United Kingdom

Corresponding Author:

Adrian Gradinar, BSc, PhD

Lancaster Institute for the Contemporary Arts

Lancaster University

The LICA Building

Lancaster,

United Kingdom

Phone: 44 1524510302

Email: a.gradinar@lancaster.ac.uk

Abstract

Background: When caring for patients with chronic conditions such as chronic obstructive pulmonary disease (COPD), health care professionals (HCPs) rely on multiple data sources to make decisions. Collating and visualizing these data, for example, on clinical dashboards, holds the potential to support timely and informed decision-making. Most studies on data-supported decision-making (DSDM) technologies for health care have focused on their technical feasibility or quantitative effectiveness. Although these studies are an important contribution to the literature, they do not further our limited understanding of how HCPs engage with these technologies and how they can be designed to support specific contexts of use. To advance our knowledge in this area, we must work with HCPs to explore this space and the real-world complexities of health care work and service structures.

Objective: This study aimed to qualitatively explore how DSDM technologies could support HCPs in their decision-making regarding COPD care. We created a scenario-based research tool called *Respire*, which visualizes HCPs' data needs about their patients with COPD and services. We used *Respire* with HCPs to uncover rich and nuanced findings about human-data interaction in this context, focusing on the real-world challenges that HCPs face when carrying out their work and making decisions.

Methods: We engaged 9 respiratory HCPs from 2 collaborating health care organizations to design *Respire*. We then used *Respire* as a tool to investigate human-data interaction in the context of decision-making about COPD care. The study followed a co-design approach that had 3 stages and spanned 2 years. The first stage involved 5 workshops with HCPs to identify data interaction scenarios that would support their work. The second stage involved creating *Respire*, an interactive scenario-based web app that visualizes HCPs' data needs, incorporating feedback from HCPs. The final stage involved 11 one-to-one sessions with HCPs to use *Respire*, focusing on how they envisaged that it could support their work and decisions about care.

Results: We found that HCPs trust data differently depending on where it came from and who recorded it, sporadic and subjective data generated by patients have value but create challenges for decision-making, and HCPs require support in interpreting and responding to new data and its use cases.

Conclusions: Our study uncovered important lessons for the design of DSDM technologies to support health care contexts. We show that although DSDM technologies have the potential to support patient care and health care delivery, important sociotechnical and human-data interaction challenges influence the design and deployment of these technologies. Exploring these considerations during the design process can ensure that DSDM technologies are designed with a holistic view of how decision-making and engagement with data occur in health care contexts.

(*JMIR Hum Factors* 2022;9(2):e32456) doi:[10.2196/32456](https://doi.org/10.2196/32456)

KEYWORDS

data-supported decision-making; health care professionals; respiratory care; scenario-based design; clinical decision-making; decision support; COPD; respiratory conditions; digital health; user-centered design; health technologies

Introduction

Background

Chronic obstructive pulmonary disease (COPD) is one of the most common chronic respiratory conditions in the world [1]. COPD typically arises from long-term exposure to airway irritants, such as cigarette smoke or air pollution [2-5]. It causes nonreversible chronic obstruction of the airways, resulting in breathlessness, fatigue, and frequent chest infections [6-8]. These symptoms can make it difficult to engage in daily activities, such as leaving home, socializing, and getting dressed [9,10]. Exposure to respiratory infections, physical exertion, smoke inhalation, and environmental factors such as air pollution can worsen symptoms [4,11-13]. COPD is a considerable challenge for millions of people and many health care services around the world [14,15]. It is estimated that ≥ 200 million people have COPD worldwide [16], with approximately 16 million people in the United States and 1.2 million in the United Kingdom [16,17]. In the United Kingdom specifically, COPD generates $\geq 140,000$ hospital admissions annually, with 97% of these admissions being for emergency care [17,18].

When caring for patients with COPD, health care professionals (HCPs) must make timely and informed decisions to treat patients effectively. Clinical decision-making is a complex process that involves using medical knowledge to make decisions about care [19]. Making informed clinical decisions about patients with chronic conditions such as COPD requires quick access to a range of information about the patient and their medical history [14,20,21]. Insight into symptoms, quality of life, medications, past interventions, and results of recent clinical tests can add important context to inform decisions [21,22]. For example, by knowing the frequency and context of a patient's respiratory exacerbations (ie, flare-ups of their COPD), HCPs can suggest more personalized interventions that may be more effective.

However, data relevant for chronic condition care are heterogeneous and often buried across paper notes or electronic records or held by other HCPs involved in the patient's care [14,22-24]. Effectively collating and visualizing data about a patient's chronic condition has the potential to support timely and informed care decisions [23-25]. This presents an opportunity to explore how digital technology can be designed to provide timely data-driven support for HCPs. Digital technology, which provides data that support decision-making, is termed data-supported decision-making (DSDM) technology. Designing DSDM technologies to support demanding health care contexts requires us to work closely with HCPs to explore their needs and expectations. An appreciation of the broader complexities of health care work is also needed [23].

In response, our research aimed to actively engage HCPs in considering how DSDM technologies could support clinical decision-making in the context of COPD care. Through extensive engagement with HCPs, we identified a set of data

interaction scenarios relevant to their practice. We then created an interactive web application as a tool to visualize these scenarios and facilitate discussion about how DSDM technologies might support their work.

DSDM Technologies in Health Care

Electronic health records [26], dashboards [27,28], and clinical decision support systems [29,30] are types of DSDM technologies used across health care. They present pertinent information to inform clinical decision-making. Dashboards are a prominent form of DSDM technology that can improve patient care [27,28,31-34]. Dashboards aggregate and visualize data in ways that produce insights to users. For instance, to support users to increase the number of patients undergoing health screening [31], identify possible high-risk medication prescribing scenarios [32], track in-patients in mental health wards [33], and effectively use patient-reported outcome data for cancer care [34].

While reviewing the literature on how DSDM technologies are designed and used in health care, we found that many studies focused on measuring the clinical effectiveness or quantified outcomes achieved using the technology [25,28,29,31,32,35,36]. Although these studies are crucial for establishing the quantitative impact of DSDM technologies on health care, they do not document the design process or provide detailed user insights about the technology. This knowledge is crucial to inform how DSDM technologies should be designed for real-world contexts from a human-centered perspective [37-40]. Collaborating with HCPs during the design process can unearth important sociotechnical and human-data interaction considerations required to build successful technologies [41-43]. Sociotechnical considerations investigate the design and implementation of systems based on technical and social dimensions [44]. Human-data interaction investigates how people interact with, interpret, and understand data [45,46]. Legibility, agency, and negotiability are key human-data interaction challenges [46]. Legibility refers to making data and algorithms transparent and comprehensible. Agency is the capacity to act on data and data implications. Negotiability relates to re-evaluating decisions about data and data processing as contexts change.

We found that a small number of studies have investigated the challenges involved in designing DSDM technologies for specific clinical contexts [47-49]. Bardram and Nørskov [48] and Sarcevic et al [47,49] took a user-centered approach to inform the design of context-aware dashboards for high-risk settings. Their prototypes focused on patient safety in operating theaters [48] and trauma resuscitation [47,49] and were evaluated with staff during a simulated clinical scenario. The researchers then revised how the data were presented to effectively support decision-making in these contexts, for instance, supporting dynamic information visualization in the fast-paced setting [48] and excluding audio feedback that could startle patients [49]. The nuances of clinical work were

understood by engaging hospital staff during the design process, highlighting the value of partnering with end users when designing DSDM technologies [44,50]. However, there is more to learn about designing DSDM technologies outside the specific use case of high-risk settings [47-49]. Crucially, we need to explore wider everyday data interactions to inform the design of DSDM technologies in health care settings. This can enhance our understanding of the possibilities for DSDM technologies in health care.

Scenario-Based Design in Health Care

Scenario-based design has been used in previous studies to evaluate health care technologies [43,48,49,51,52]. Scenarios, which are task-driven descriptions of work instances, focus on how a system can support human activities [51,53]. They are effective for the qualitative systematic evaluation of usability, suitability, and user experience of a technology or prototype [51,53]. This involves users completing tasks presented as scenarios on the proposed tool, presented as scenarios, and evaluating their experience.

Scenarios are particularly effective in eliciting detailed feedback from users without deploying a full system in clinical practice [52]. For example, Bardram [51] used scenarios to approach the redesign of an information system used in hospitals. Using scenarios allowed the hospitals' existing activities to remain central to the design and evaluation, helping to focus on how the system could support both current and future activities. Scenarios are a creative thinking tool for envisaging how systems can support work and how it is organized [51]. Given our desire to capture rich details about how DSDM technologies could support HCPs with COPD care, scenario-based design was an appropriate method for this study.

Study Aims

This study explored how DSDM technologies could support HCPs in their decision-making regarding COPD care. We achieved this by presenting an exploration of a scenario-based research tool called *Respire*. *Respire* is an interactive web app that presents HCPs with data interaction scenarios to support their decision-making about their patients with COPD and service. We designed *Respire* with input from 9 respiratory HCPs over 2 years and subsequently explored the output with 11 respiratory HCPs (9 of which were involved in the design process).

Our findings uncover the challenges faced when HCPs interact with health care data in context. From this, we reveal novel insights and lessons regarding the design of DSDM technologies to support the real-world complexities of clinical decision-making. Our paper makes three main contributions: (1) we provide insights into how DSDM technologies can support respiratory care by exploring HCPs' data needs; (2) we uncover key barriers that impact HCPs' engagement with data for decision-making; and (3) we provide novel and translatable [50] design implications that inform the creation of future DSDM technologies for health care.

Study Structure

This study was divided into 3 stages. The first stage explores HCPs' data needs related to COPD care, with a view to understanding how DSDM technology could support these requirements. The second stage involves the selection of key data requirements identified from the first stage and developing them into digital *data interaction scenarios* (presented in *Respire*). The third stage explores *Respire* with HCPs to understand how each data scenario could support their decision-making regarding COPD care.

Methods

Overview

This was a co-design study involving HCPs from 2 collaborating National Health Service organizations in North West England. Co-design involves embedding users in the design process, which is appropriate, given our desire to explore HCPs' needs and experiences at each stage in detail [54]. The first organization we worked with, *the hospital*, has a respiratory ward with patients with COPD under the care of specialists. The specialists also visit respiratory outpatients in clinics, including those recently discharged from the hospital after an exacerbation. The second organization, *community care*, provides services that enable patients with COPD to manage their condition in the community. Services include routine clinics to assess a patient's condition and management, pulmonary rehabilitation classes that use exercise and education to improve self-management [55], and home support services in which on-call specialized nurses support acute patients in their homes [56]. Patients are referred to community care by their general practitioner (GP) if they require advanced support or by the hospital to help stabilize their management after a hospitalization.

We have previously worked with HCPs from both organizations to explore their challenges with lack of access to data and effective visualizations for COPD care [24]. In a study by Tendedez et al [24], we found that (1) HCPs used multiple clinical systems that were inflexible, (2) existing data lacked required detail and quality, and (3) HCPs rarely had time for extensive training on clinical systems and needed intuitive user interfaces. Following Tendedez et al [24], we worked with them to explore how DSDM technology could support their needs by creating *Respire*. *Respire* is a web app designed to effectively visualize COPD data that are routinely collected across both organizations (contained in digital systems or paper notes). Crucially, it also aimed to visualize data that are not yet available in clinical practice to envision how decision-making could be supported in the future. The current and future data needs were identified through an iterative process.

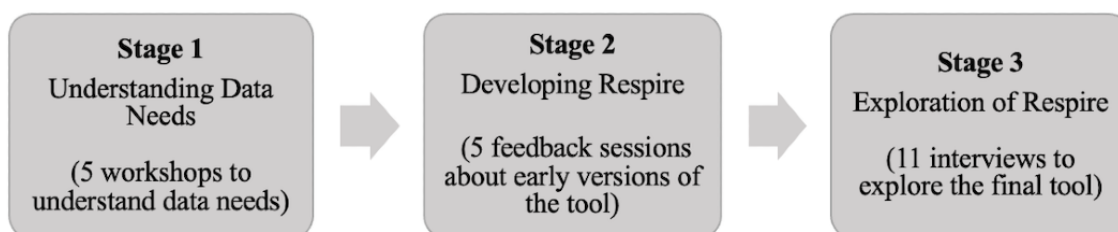
We had no access to the organization's data during this study. Therefore, *Respire* used test data sets that were created by the research team and based on the data requirements elicited during the study. A researcher experienced in large-scale hospital data advised on the content and structure of the test data sets. For instance, they advised on (1) the typical data fields that hospitals collect for patients with COPD and (2) the range of data within those fields. For scenarios that displayed medical data, we used

web-based resources regarding medical readings to inform clinically realistic test values [57]. The researcher then checked that the test data sets we had produced were realistic for the purposes of the activity. The test data were created solely to populate Respire for this study; they were not intended to be

used beyond this purpose. We stored the data in a MySQL database, which was read by Respire via a custom REST (Representational State Transfer) application programming interface.

An outline of the study methodology is presented in Figure 1.

Figure 1. The 3 stages of the study methods.



Participants

This study involved 11 HCPs across both organizations (6 from community care). Each participant (except for C6 and H5) was involved in the design process. Snowball sampling was used to

recruit participants, with HCPs informing their colleagues about the project [58]. Details of the participants is presented in Table 1, showing their experience in their current role and using clinical information systems.

Table 1. Details of study participants.

Participant identifier ^a	Role	Years in current role	Experience using clinical information systems (years)
H1	COPD ^b nurse	<1	3
H2	COPD nurse	<1	17
H3 ^c	Respiratory consultant	5	18
H4 ^c	Respiratory consultant	3	2
H5 ^d	Respiratory consultant	6	9
C6 ^e	Respiratory service manager	2	25
C7 ^c	Lead COPD nurse	2	13
C8	COPD nurse	14	21
C9 ^c	Lead physiotherapist	6	10
C10	Assistant practitioner	7	10
C11	COPD nurse	12	12

^aIdentifiers prefixed with H are from the hospital and C are from community care.

^bCOPD: chronic obstructive pulmonary disease.

^cStudy champions were contact points that helped to coordinate research sessions.

^dH5 was invited to participate in the study by H3 after Respire was designed.

^eC6 was involved in early discussions but was unavailable to participate in the design process.

Data Analysis

This qualitative study followed an interpretivist approach that emphasizes the social construction of individuals' knowledge based on their lived experiences [59]. This approach was appropriate because of the exploratory nature of this study, which focused on HCPs' experiences providing COPD care. We used 2 techniques for data analysis.

Stages 1 and 2 used content analysis to determine the presence and frequency of specific themes within discussions [60]. Content analysis was selected because of the volume and nature of the data collected. We carefully read the transcripts and

assigned codes to references to specific types of data, reasons the data are needed, and comments about data visualization. Stage 3 used inductive thematic analysis to analyze the interview data. We used bottom-up open coding to assign codes to the data at the sentence level based on what the data described. We then grouped those codes to create broader themes that described the entire data set [61]. This analytical approach was chosen given our desire to be more exploratory in stage 3, focusing on capturing the nuances of interacting with Respire.

Understanding Data Needs (Stage 1)

The first stage involved 5 workshops to explore the data needs for Respire. There were 2 separate workshops with the hospital (H1, H3, and H4) and community HCPs (C7, C8, C9, C10, and C11) each before uniting in the final workshop. Sessions were organized to suit HCPs' availability and lasted between 30 and 90 minutes in quiet rooms at the clinical sites. Plans for workshops 1 and 2 are provided in [Multimedia Appendix 1](#). We analyzed workshop transcripts after each session.

The first workshop aimed to understand the data needs and develop a shared language between HCPs and researchers. The discussions focused on their patients with COPD and services. We asked, "What data do you want to see about your COPD patients/service?" and "How would you want to interact with that?" In addition to verbal discussions about data needs, to stimulate discussion, HCPs created basic sketches of how data might be visualized. Basic sketching was used as a visual communication tool, enabling researchers to understand HCPs' mental model of how the data might be tangibly presented. After the sessions, we used the sketches to create wireframes of basic user interfaces and to complement the data analysis. We provided the wireframes in later workshops as stimulus materials.

Subsequent workshops explored the data needs in detail using the wireframes to structure discussions. The HCPs revised the way the data were visualized on the wireframes and refined the included data. For instance, they supplemented the tables with graphs and removed data that they no longer saw as a priority on reflection. We updated the wireframes after the session.

The final workshop gave HCPs from both organizations the opportunity to discuss one another's wireframes. Two business intelligence staff members, who had been involved in earlier stages of the project [24], attended to share knowledge about the general existence of the included data within both organizations.

Developing Respire (Stage 2)

We reviewed the data needs captured in stage 1 and chose 5 key use cases to expand into data interaction scenarios for Respire (Figures 2-10). We decided to focus on the needs common to both organizations, as these appeared to be most impactful. For example, we created a scenario based on viewing a patient's spirometry test result history, as this was an unmet need voiced by both organizations. Spirometry tests require the patient to blow into a device used to diagnose and monitor respiratory conditions. Both organizations discussed these data as being a prominent pain point in practice; thus, it was an important scenario to explore further.

Figure 2. Scenario 1: respiratory ward overview (annotated). This view lists the patients with chronic obstructive pulmonary disease (COPD) in hospital for a COPD-related reason. (A) lists the ward that the patient is on, (B) details each patient's current length of stay in days, (C) details the number of COPD hospital admissions each patient has had in the past 12 months.

Respiratory Ward Overview						
Overall		Hospital Site 1			Hospital Site 2	
Primary COPD admissions	11	Total Patients on COPD marker			55	
Secondary COPD admissions	9					
COPD primary admission reason				COPD secondary admission reason		
Name	NHS no.	Hospital no.	Bed no.	Ward	Current length of stay	Prev COPD admissions < 12 mo.
patient0	12345	10293	1	Respiratory	2	1
patient1	54321	29384	2	Respiratory	4	0
patient2	23456	39848	3	Cardiac	2	1
patient3	65432	12345	4	Respiratory	2	4
patient4	34567	43543	5	Respiratory	5	0
patient10	67890	89435	11	(A) atory	(B)	(C)

Figure 3. Scenario 2: Admissions and Exacerbation Reports (annotated) showing overall hospital admissions. (A) shows whether the admission was chronic obstructive pulmonary disease (COPD) related or not (B) shows COPD-related hospital admissions split between patients previously known to have COPD and those newly diagnosed because of the admission.

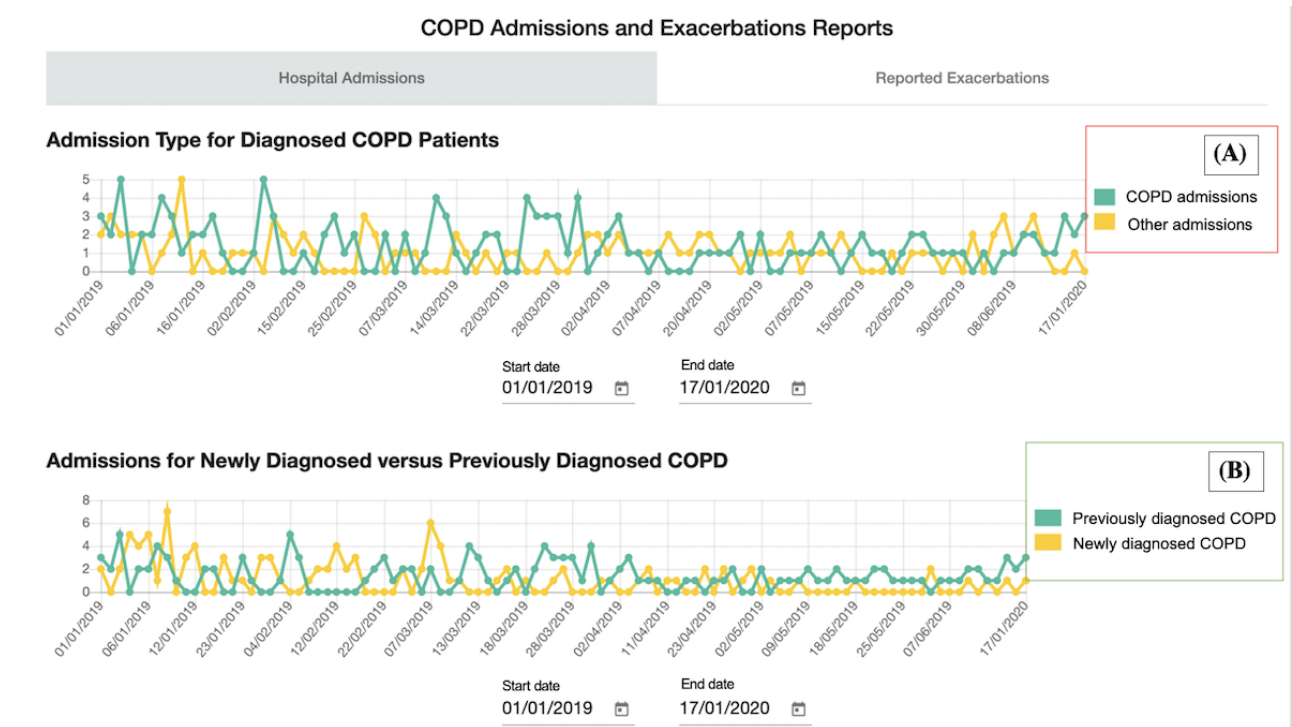


Figure 4. Scenario 2: Admissions and Exacerbation Reports (annotated) with a tab for overall reported infective exacerbations and chronic obstructive pulmonary disease (COPD) hospital admissions. (A) shows which service reported the exacerbation. GP: general practitioner.

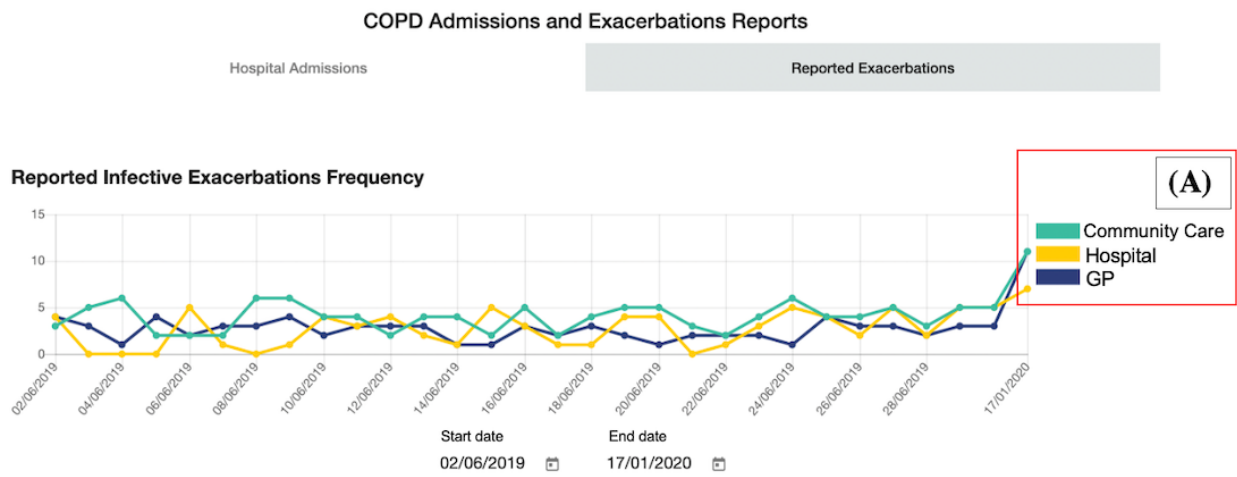


Figure 5. Scenario 3: patient-generated data overview (annotated). (A) Patients’ latest symptom status (green indicates asymptomatic, amber indicates symptomatic, and red indicates severe symptoms); (B) the traffic light system depicting patients’ 7-day status (gray indicates no data have been entered by the patient).

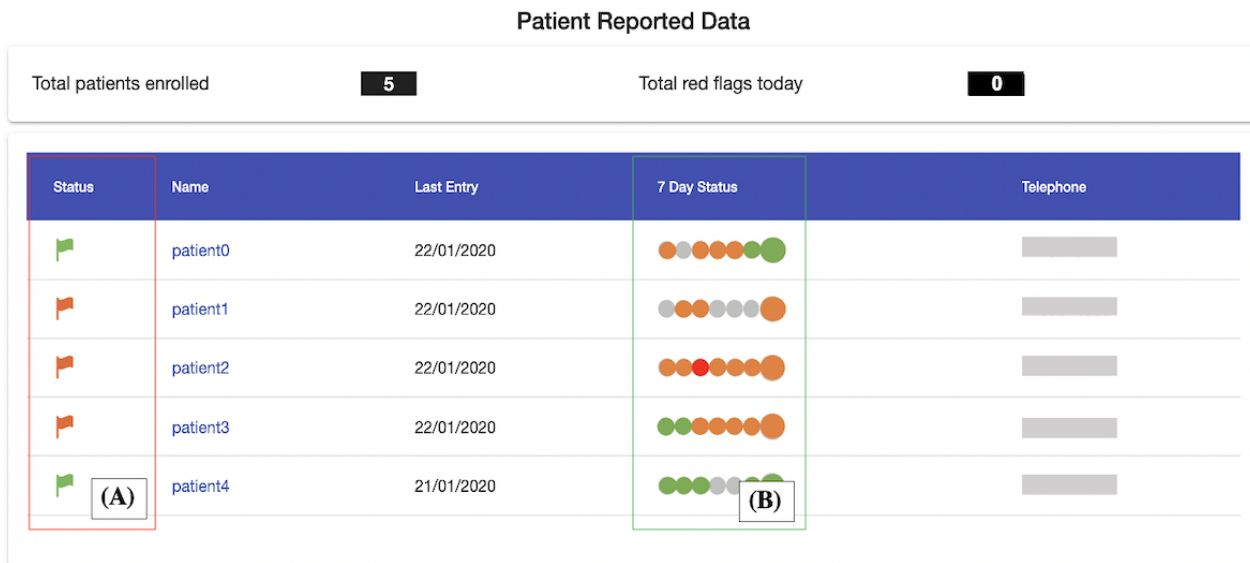


Figure 6. Scenario 3: patient-generated data individual entries (annotated). (A) A log of a patient’s symptom entries and (B) the logged entries in graph format.



Figure 7. Scenario 4: example patient’s exacerbation history. The table shows a history of a patient’s clinically reported chronic obstructive pulmonary disease (COPD) exacerbations; (A) shows which service has managed each exacerbation. GP: general practitioner.

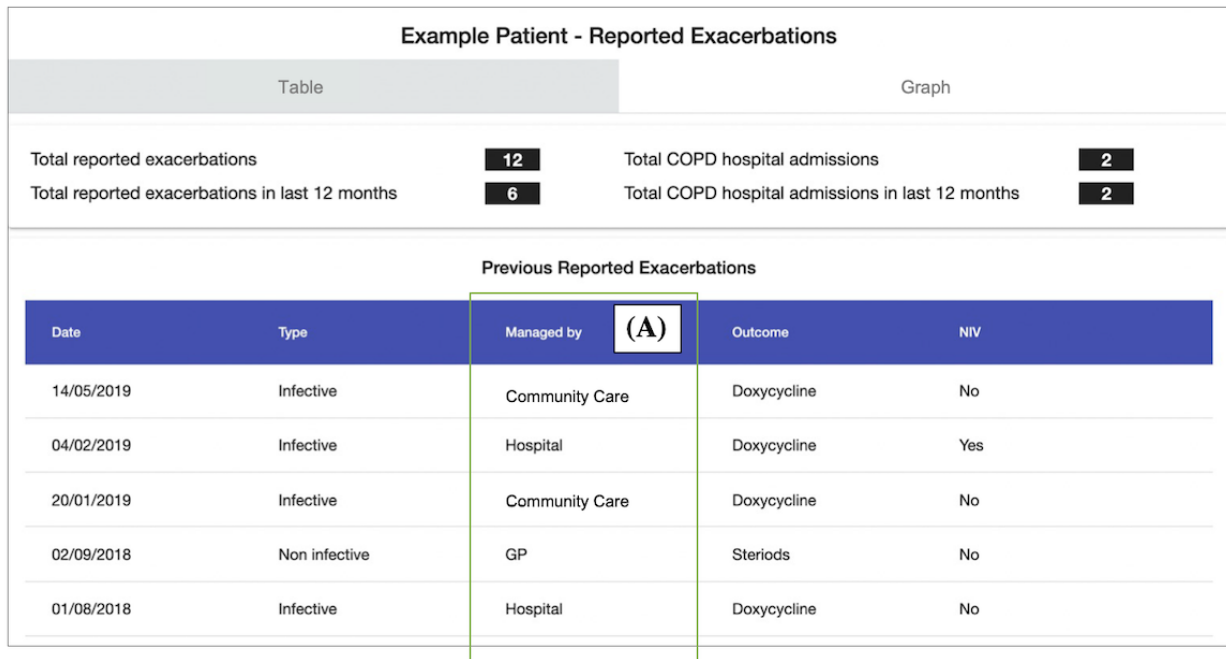


Figure 8. Scenario 4: example patient’s exacerbation history. The graph represents the frequency of the clinically reported exacerbations of a patient with chronic obstructive pulmonary disease over time; (A) shows how health care professionals can filter by year; (B) shows which service has reported the exacerbation. GP: general practitioner.

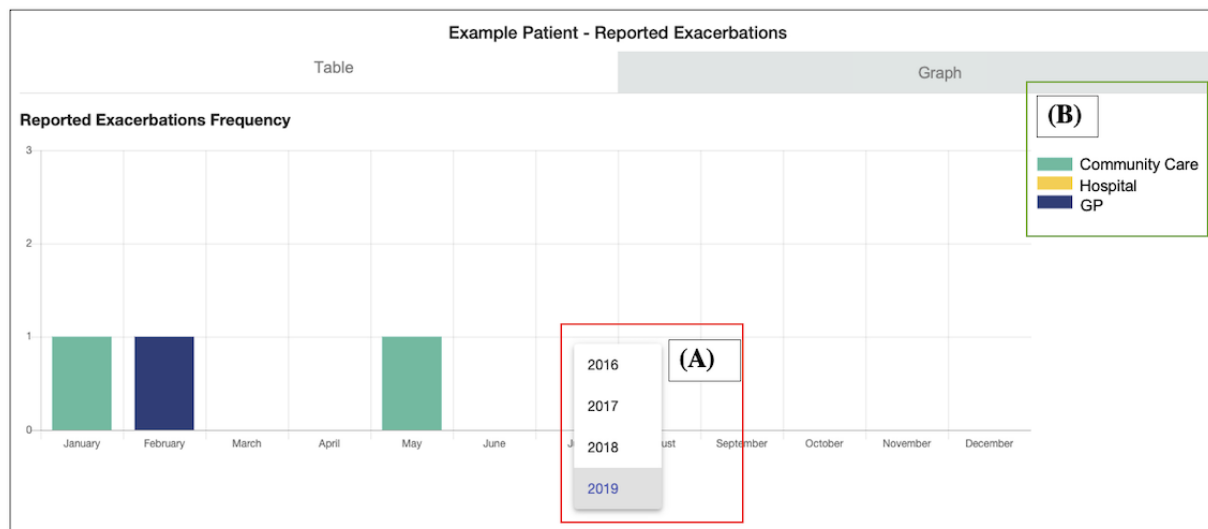


Figure 9. Scenario 5: example patient’s spirometry results; (A) the spirometry trace for the test result; (B) where the test was taken. FEV: forced expiratory volume; FVC: forced vital capacity; RV: residual volume.

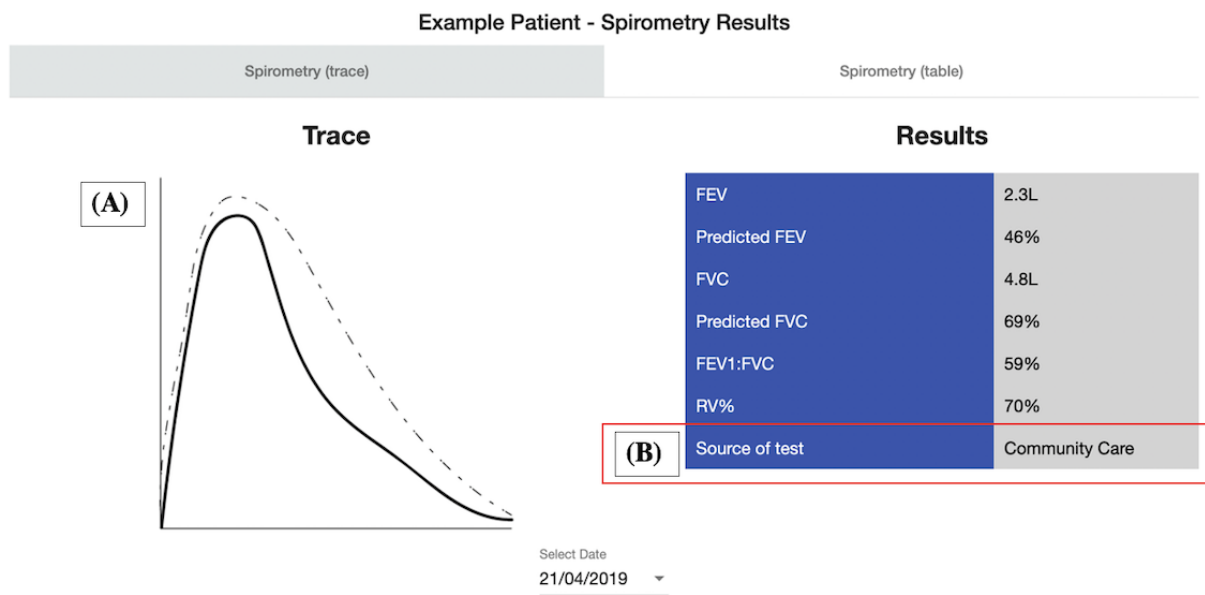


Figure 10. Scenario 5: example patient's spirometry results. The table shows a breakdown of all the spirometry test results for a patient; (A) which service the test was taken at. GP: general practitioner. FEV: forced expiratory volume; FVC: forced vital capacity; RV: residual volume.

Example Patient - Spirometry Results

Spirometry (trace)				Spirometry (table)					
Date	FEV1	Pred.	FVC	Pred.	FEV1:FVC	RV%	Source	(A)	Trace
21/04/2019	2.3L	46%	4.8L	69%	59%	70%	Community Care		View
01/11/2018	2.0L	40%	4.8L	55%	40%	70%	GP		View
14/02/2018	2.3L	50%	4.8L	72%	60%	70%	Hospital		View
03/08/2017	2.0L	42%	4.8L	53%	44%	70%	GP		View

Following this, we confirmed with H4 and C7 if the 5 chosen scenarios (Table 2) were an important focus. We then created Respire, which is an interactive digitization of the 5 data scenarios. Respire is a web app populated with test data to support dynamic interaction. During development, we met with available HCPs (H2, H3, H4, C7, C8, and C9) for feedback on

the early versions. For example, ensuring that the wording of headings and the test data made sense for each scenario. Table 2 details the 5 shortlisted scenarios, including a description of the scenario and where the data contained within the scenario would be reported from. A snapshot of each scenario on Respire is shown in Figures 2-10.

Table 2. The 5 shortlisted data scenarios.

Number	Scenario name	Scenario description	Data reported by
1	Respiratory Ward Overview	List of in-patients with COPD ^a , the ward they are on, length of stay, and their number of previous COPD-related admissions	Hospital
2	Admissions and Exacerbation Reports	Reports on population-level COPD hospital admissions and exacerbations. Live and historical data can be viewed	Hospital, community care, and GP ^b practices
3	Patient-Generated Data Overview	View of patients using a mobile app to self-monitor their cough, breathlessness, sputum production and color, and actions in response to symptoms	Patients with COPD
4	Example Patient's Exacerbation History	Overview of clinically reported exacerbations of a patient with COPD and which service reported them	Hospital, community care, and GP practices
5	Example Patient's Spirometry Results	A full history of spirometry test results of a patient with COPD and which service the test was taken at	Hospital, community care, and GP practices

^aCOPD: chronic obstructive pulmonary disease.

^bGP: general practitioner.

Exploration of Respire (Stage 3)

The final stage explored Respire with 11 HCPs (Table 1). We gathered feedback about their interaction with the scenarios, exploring how they might support their decision-making regarding COPD care. A total of 11 one-to-one sessions were held in quiet rooms at both clinical sites, lasting between 60 and 90 minutes. HCPs were tasked with walking through each scenario and imagining that they had access to it in practice. We asked HCPs to interact with each scenario freely by exploring different tabs, reviewing and interacting with the visualizations, and examining the (test) data. They were asked to think aloud [62]. The semistructured session plan is contained in Multimedia Appendix 2. During the sessions, we asked, "How could the data presented to you in this format influence your decision-making?" "Are there any challenges that you could envisage when using this scenario?" "Who do you think needs to be involved in the collection and maintenance of this data to ensure it is useful?" These questions would provide insight into how HCPs envisage digital tools, such as Respire, might be used in practice.

After reviewing each scenario, HCPs rated (on a 7-point Likert scale, with 1 representing *strongly disagree* and 7 representing *strongly agree*) how realistic they perceived the digital data scenario to be ("the scenario responds in a way that you would expect when using a system to complete this task") and its relevance to their job ("this scenario is something you would use in your role"). The former was asked to understand if the scenarios were presented realistically to inform them how well

they could engage with them and respond to questions. The latter was asked to understand whether there were scenarios that were more relevant to the responsibilities of some HCPs compared with others. Both these responses would add further framing to the discussion.

To conclude the sessions, HCPs ranked the 5 scenarios against each other in order of usefulness (with a score of 1 being the most useful). This would help to discuss the respective strengths and weaknesses of each scenario in context with another.

Ethical Considerations

This study received ethics approval from both Lancaster University's Faculty of Science and Technology Ethics Committee and the Health Research Authority (reference: 17/HRA/3092). All participants were required to read an information sheet and sign an informed consent form before participation. All sessions were audio-recorded, with full permission from the participants.

Results

Understanding Data Needs (Stage 1)

The data requirements shared by both the hospital and community care were identified from the stage 1 workshop and are summarized in Textbox 1. They largely focused on (1) understanding the severity of a patient's condition and (2) managing the demands of both health care services. A full list of the data requirements captured from the stage 1 workshop is provided in Multimedia Appendix 3.

Textbox 1. A summary of the hospital and community care’s shared data requirements with direct quotes from participants during stage 1.

Patient’s spirometry result history

- “To *know* if it’s definitely COPD. Then if it is, then what was it [the result] before, does it mean that it’s getting worse [now]?” [H4]
- “The shape of the curve [trace]...will tell you potentially a bit more about their airways. We generally just have the numbers.” [C9]

Patient’s previous chronic obstructive pulmonary disease (COPD) hospital admissions and exacerbations

- “When you see patients from admission to admission you might not necessarily join everything together.” [H3]
- “In a certain timeframe how often have they been admitted? Three plus exacerbations, then I would consider that is a suitable patient for us [to manage as opposed to the GP].” [C9]

Patient-generated data about their COPD symptoms

- “What has the patient’s perspective been of their illness...we need to understand what the patient understands about their illness.” [H3]
- “Capturing exacerbations and deterioration earlier to avoid potential hospital admissions and potential deterioration.” [C9]

Patient’s respiratory medications and breathlessness rating

- “[It] impacts upon how we might manage them...have they been getting more breathless...have their treatments changed?” [H3]
- “If they’d been given a rescue pack of antibiotics and steroids [to take at the onset of exacerbations at home]...[and] to know if they’ve had, say, 6 antibiotics in the last 3 months.” [C8]

Respiratory interventions a patient has had

- “[You say] this patient has had 2-3 admissions needing non-invasive ventilation (NIV), have you thought about domiciliary NIV? Or they’ve not done pulmonary rehab in over a year...could you do that?” [H4]
- “You could see what’s been offered or if they’ve been referred [for interventions] and declined.” [C9]

Live list of COPD-related admissions at the hospital

- “How many have been there [on the ward] since last week that we need to target first so we can facilitate their discharge?” [H4]
- “[Currently the system] brings up a list of COPD patients...it won’t say whether the particular admission is because of their COPD.” [H3]
- “We actually need to be targeting some of these [admitted] patients that aren’t accessing us [in community care].” [C7]

Exploration of Respire (Stage 3)

The following sections outline the findings from the stage 3 interviews, including Likert questionnaires, scenario ranking, and qualitative feedback.

Quantitative Scenario Feedback

Table 3 presents the results of the Likert questionnaires, showing the mode of participants’ ratings across each scenario and the frequency of the mode. For the scenario realism and relevance scores, 7 indicated *strongly agree*, 4 indicated *neither agree nor disagree*, and 1 indicated *strongly disagree*.

Table 3. Results from the stage 3 Likert questionnaires.

Scenario number and scenario	Realism score ^a		Relevance score ^b	
	Mode ^c	Frequency of mode	Mode ^c	Frequency of mode
1 Respiratory Ward Overview	7	7	7	7
2 Admissions and Exacerbation Reports	7	4	1	4
3 Patient-Generated Data Overview	7	4	6	5
4 Example Patient’s Exacerbation History	7	7	7	7
5 Example Patient’s Spirometry History	7	5	7	7

^a“The scenario responds in a way that you would expect when using a system to complete this task.”

^b“This scenario is something you would use in your role.”

^c7 indicates *strongly agree*, 4 indicates *neither agree nor disagree*, and 1 represents *strongly disagree*.

The most common realism rating was *strongly agree* (score=7) across all scenarios. Scenarios that commonly received the highest relevance ratings were scenario 1 (Respiratory Ward

Overview), scenario 4 (Example Patient’s Exacerbation History), and scenario 5 (Example Patient’s Spirometry History). Scenario 2 was commonly rated the lowest for relevance (Admissions

and Exacerbation Reports). Usefulness scores are presented alongside the qualitative findings in the following sections for context.

Qualitative Scenario Feedback

Scenario 1: Respiratory Ward Overview

Scenario 1 was ranked as the most useful (ranked first place by 6 participants), with the main benefit being that HCPs could quickly identify patients who required support. Hospital HCPs believed that the length-of-stay indicator would help identify patients who have been in hospital the longest to prioritize during ward rounds. It would also help assign senior staff to patients with the longest stays, as these patients are likely to have complex health needs. Similarly, viewing the number of each patient's previous COPD hospital admissions would allow them to be supported in specific ways. For example, patients without previous admissions may benefit from education on managing their condition. Patients with many previous admissions may require end-of-life care. Community care HCPs felt that the scenario could help identify patients who were in the hospital for their COPD to offer support on discharge. Currently, to achieve this, they must "trawl" [C7] through a list of discharged patients known to have COPD without easily seeing why the patient had been in hospital.

Knowing the data source that would populate scenario 1 was key for HCPs to consider its drawbacks when making their decisions. HCPs explained that the 2 existing data sources that could show which patients with COPD were in hospital had inaccuracies. The first data source was a list of patients with COPD from their data flag system. This system flags patients diagnosed with COPD by local GP practices, hospitals, or community care. However, it is not a "true list" [C7] as (1) patients on the list sometimes "have other respiratory conditions" [C11] and are incorrectly diagnosed with COPD and (2) the system is "not utilized very well" [C7] as flagging patients is a manual process and some patients "probably slip through the net" [C7]. The second data source was the hospital's clinical coding department. The initial coding of a patient's hospital admission reason is done by emergency department staff, who are usually "generalists" [C6], and their working diagnosis does not always reflect the final reason for admission. In addition, "very umbrella type codes" [C6] within current classification systems (such as International Classification of Diseases, Tenth Revision [63]) indicate that multiple codes can describe a single hospital admission. For example, COPD may be coded as either *COPD* or *breathlessness*. This means that admissions coded as *breathlessness* could have been missed from the data set that populates scenario 1.

Scenario 2: Admissions and Exacerbation Reports

Scenario 2 was ranked the fourth most useful scenario (ranked fourth place by 5 participants), with the main benefit around supporting service planning. For example, to see "where people are referring themselves [when they are unwell]...that first presentation [of symptoms]" [C9], so that the service can

identify where they may need extra resources. Forecasting hospital admissions was another way to plan services based on the data as "GP [appointment] spikes normally occur slightly before admission spikes, so if there is starting to be a GP spike then you can follow the trend" [H5]. HCPs discussed how this scenario would be checked on a "monthly" [H4] basis.

However, a key challenge for scenario 2 was the perceived lack of a consistent understanding of COPD exacerbations across hospitals, community care, and GP practices. The HCPs strongly believed this affected the quality and reliability of reporting, as exacerbations are labeled "too easily" [C9]. H4 described this in detail:

It's easy to label them [patients] as having an exacerbation and give them a little bit of steroids and a little bit of antibiotics...it comes back to how much do you trust the person who is saying they have taken it seriously and taken it to say this is an actual exacerbation?

Which was echoed by C11:

I also do feel like from a professional side that medics are like "well we'll give you this [treatment for an exacerbation] because it'll move you on through and out the system"...I do think there's a bit of discrepancy [about what exacerbations are]

The differences in exacerbation reporting were thought to exist because not all HCPs who see patients with COPD are specialists in COPD. The hospital's and community care's specialism in COPD makes their identification of exacerbations more reliable, compared with GP practices and emergency department staff who generally do not have COPD "expertise" (H1). Furthermore, HCPs without COPD expertise could assume the patient "knows their condition best" (H1) when approached about a suspected exacerbation and thus provide treatment for an exacerbation.

Scenario 3: Patient-Generated Data Overview

The usefulness ranking for scenario 3 was bimodal, ranked least useful by 3 participants and second most useful by 3 participants. This scenario was seen as valuable for understanding the overall patient experience of living with COPD. HCPs felt that it could be used as a tool to educate patients on their condition. In the clinic, the data could be "an entry to a conversation" [C6] about what actions the patient could take when experiencing certain symptoms. For example, when looking at the (test) data, H4 saw a patient in contact with their health care team despite reporting no symptoms (Figure 11). They felt the patient could have anxiety about their COPD and need "assurance," with discussions focusing on how the patient could help themselves when they feel anxious. C6 discussed using the data similarly to suggest to the patient "some breathing techniques to help," so they could distinguish breathlessness caused by anxiety versus an exacerbation. This is important, as anxiety can influence feelings of breathlessness, which patients might not differentiate from a respiratory exacerbation [64-66].

Figure 11. Scenario 3: patient-generated data overview showing an example patient's symptom log where they had no symptoms but contacted their health care team.

patient2 - patient generated data

Self-reported entries (table)			Self-reported entries (graph)			
Date	Day Type	action	Breathlessness	Cough	Increased sputum	Sputum Colour
22/01/2020	No symptoms	I made contact with my health team	None	No cough	No	Good
21/01/2020	No symptoms	I took more reliever	None	No cough	No	Good
20/01/2020	flare up	I took steroid tablets	Breathless	Increased	Moderate	Bad
19/01/2020	flare up	I took steroid tablets	Breathless	Increased	Moderate	Bad

However, identifying the right patient group for self-monitoring is crucial for scenario 3 to be “constructive” [C6]. For patients stable in their management, self-monitoring may be “medicalizing their condition” [C6] and be a reminder “that they are ill” [H5]. However, HCPs felt that acutely ill patients with several previous hospital admissions could benefit from self-monitoring. Newly diagnosed patients could also self-monitor to become familiar with their symptoms. HCPs talked about patients using this platform for a specified period of time for encouragement rather than indefinitely. For example, C9 suggested that patients “could be put on it for a month and monitored by the community care team” to combat 30-day hospital readmissions that occur with COPD [67]. In addition to identifying the right patient group, it was also important for patients to input the required data. Consistent data entry would provide a “true overall reflection” [C10] of a patient's condition, to be “sure about the day-to-day changes” [C8]. Others felt that sporadic data entries could be acceptable as long as the data are entered when the patient is symptomatic. For example, H5 discussed how:

if you have loads of grey [no input] and then three red [severe symptoms], you know you need to phone them...but there will also be patients who just don't put data in until they are unwell. What you don't want is patients who put greens [asymptomatic] but don't put the red.

Some HCPs discussed how, in certain contexts, asymptomatic days could be “hidden” [C9] from view as “there's no need to worry about them” [C9]. Despite this, HCPs stressed the importance of recognizing a patient's symptom-free period, which could be challenging to reinforce if there is a focus on recording only when symptomatic.

Following the need to receive enough data to support decisions, HCPs felt concerned about what missing data might mean and the resources required to investigate. C7 felt missing data could indicate that a patient may “potentially be at home isolated or be dead.” C6 described the likely process of investigating missing data:

You ring them [patient] up and they don't answer, really common...you maybe try every day for a month. At some point, you are gonna have to send them a letter or do something else” which eventually leads

to “generating a whole heap of work...you'll get through to the patient who will say ‘ah yeah I didn't bother, I'm not bothered about it anymore’.

There were also concerns about investigating the data that had been received. HCPs raised a key issue about feeling responsible for the data that the patient reports. C6 highlighted that remote setups are “implying somebody is monitoring it [the data]” and patients “may become dependent” on the idea that the HCP is continually “looking at that [data] and acting.” However, HCPs lack time and resources to instantly detect health concerns from the data. C6 was further concerned that scenario 3 could encourage patients to take less responsibility for their health concerns as “in reality it means a clinician managing them...they aren't fully self-managing under this scenario.”

Finally, trust concerns were raised regarding the self-assessed symptom data versus quantitative physiological data. Although self-assessed data were valuable to understand quality of life and patient experience, it was not reliable as “some patients will overreport their symptoms and some will under-report” [H6]. For instance, breathlessness and fatigue have a “huge psychological element” [H2] that can influence how patients perceive their symptom severity. The benefit of physiological data is “you've got a guidance that you can say ‘that's acceptable, that's not acceptable’” [C8]. However, HCPs felt self-assessed data, paired with physiological readings, were best for identifying what support to offer patients:

if someone was telling me they feel absolutely awful...but actually their physiological parameters were fine, I'd feel more reassured that perhaps they aren't clinically deteriorating, but obviously I still need to address the fact that the patient feels like they are. [H2]

Scenario 4: Example Patient's Exacerbation History

Scenario 4 was ranked jointly as the second most useful scenario (ranked second place by 5 participants), and the main benefit was a better understanding of the patient's condition journey. More specifically, “how patients' quality of life and clinical health has been affected across all sectors of care” [H5], as HCPs see patients at specific intervals depending on the concern and “what you don't see is what's been happening and how many times” [H5]. This was particularly valuable, as COPD is managed by a diverse clinical team and having “the overall

picture rather than just snippets of information” [C8] was important for effective care planning. Collating data in this way could also reveal patients who are struggling with their condition and may require a specialist referral or need “advance care planning” [H3]. Overall, HCP felt seeing past exacerbations in this way was an improvement over asking the patient about their history and shuffling through “thousands of records” [C6], with detail that is too “heavy” [H5] and “not relevant for what we [respiratory HCPs] are interested in” [C6].

Scenario 4 could also provide context for the patient’s experience of living with COPD. H4 envisioned using the scenario when engaging with patients in clinics, whereby “you sit with them to say ‘tell me what happened there’” about each exacerbation to learn about their experiences and triggers. This was seen as a valuable communication aid, as patients’ impromptu recall about their experiences “isn’t always great” [H3]. H4 added that better understanding patients’ experiences can support conversations around management:

if they are only breathless because they have seen something on the television that upset them...that has affected the way that they are feeling, but physiologically they don’t need steroids [to manage it]

Scenario 5: Example Patient’s Spirometry

Scenario 5 was ranked the third most useful scenario (ranked third place by 4 participants). Although it was seen as being “really useful” [H5] and time saving, it was perceived as less impactful to patient care than other scenarios. The main benefit was observing how a patient’s lung function may have changed over time. This was possible by comparing the spirometry result history in the table. The trace of each spirometry result alongside its numerical reading was “really important” [C11] for decision-making. This was because the trace helped to determine the “quality” [C9] of the test, it tells HCPs “how the patient performed [during the test]” [C11].

However, HCPs highlighted that the trustworthiness of spirometry test results was a key concern. HCPs felt that results from tests taken by hospital HCPs were most reliable, as not all HCPs are adequately trained to deliver spirometry tests effectively. They also felt more confident about tests taken by HCPs or services in which they had a close working relationship. C7 discussed how their close working relationship with the hospital HCPs meant they were aware of each other’s specialisms and competencies in COPD and spirometry. They described how they placed confidence in the test results from the hospital over those from GP practices:

I can see on this one (pointing to spirometry results on the screen) that this was done here [in community care], and this one at the Hospital, so you’d be more inclined to use the Hospital data as kind of reliable, that’s your reliable one, then you can probably work from that as to whether or not the others were really done properly. [C7]

In the abovementioned example, the hospital’s result was used as a baseline to judge the reliability of the rest of the results. The HCPs place different “confidence intervals” [H4] on the

data, depending on their source. This approach was observed in other HCPs: “was that [spirometry test] actually done by the hospital or community care? In which case, then it’s reliable. Otherwise, it might have been a GP” [C9]; “I definitely believe what came from the hospital over the GPs” [C8]; and “I know you’ve got who’s done the trace, so I think that gives you an idea of the reliability of it” [H5].

Discussion

Principal Findings

This study explored how DSDM technology could support COPD care. We achieved this by designing a scenario-based research tool with HCPs to understand human-data interaction for decision-making. DSDM technologies have clear potential to connect HCPs with pertinent data to inform decisions. However, we have unearthed important challenges and lessons relevant to the success of DSDM technologies in practice, which are of particular relevance to the human factors research community: (1) data recorded by HCPs may not be trusted for decision-making, (2) transparency about data sources is required to trust and understand data, (3) sporadic and subjective data generated by patients have value but create challenges for decision-making, and (4) HCPs require support to interpret and respond to new data and its use cases.

Data Recorded by HCPs May Not Be Trusted

Data were considered most trustworthy when the HCP who recorded it was perceived as an expert in assessing COPD. Previous work has shown that the source of medical information determines its adequacy for use in decision-making [68-70]. Specifically, Cicourel [68] observed how the perceived credibility of medical information was based on social and professional hierarchies. For example, they found that diagnostic information from attending physicians was rarely challenged and perceived as more objective than that of medical students [68]. In our study, data recorded by the respiratory ward staff at the hospital (perceived as highly specialized in COPD) were considered the most trustworthy, whereas data generated by GPs (perceived as less specialized in COPD) were considered the least trustworthy.

It was easier to assess if the data were trustworthy when it was produced by a familiar colleague, enabling the HCP to assess the colleague’s skills and competencies. Jirotko et al [71] described this as “biographical familiarity,” a predicate for trust. They observed how mammogram readers became familiar with the strengths and weaknesses of their colleagues, affecting how they read the mammograms produced by different centers [71]. Similarly, our study shows how a lack of biographical familiarity impacts HCPs’ engagement with data from staff with unfamiliar competencies. In contrast, the hospital and community care trusted each other’s data, as they were familiar with one another’s competencies.

Awareness of how trust impacts the use of data across departments and organizations is important and impacts how data should be displayed on DSDM technologies. Respire showed the *source* of spirometry test results and exacerbation reports, which HCPs felt were crucial contextual metadata to

emphasize. One possible way to support building trust with unfamiliar data could be through *seals of approval* or *digital badges* built into dashboard designs [72]. For example, a badge representing skill proficiency could be displayed next to entries from organizations that have had specific training in spirometry.

Transparency About Unreliable Data Sources Is Needed

Knowing which data sources were populating Respire was important for assessing any limitations when using data to make decisions. The explicit mention of a system's data sources is also important for building trust [73]. This emphasizes the need to make data sources transparent to users, addressing the human-data interaction challenges of legibility [46]. Certain data sources were perceived as unreliable, such as data from the coding department and the hospital data flag system. The unreliability of coded data has been explored in previous work [74-76], particularly the overlap of codes for a single clinical condition [74]. As the specificity of medical data is tailored to the original purpose of its collection, repurposing it requires additional details for the data to be usable in new contexts [77,78].

We argue that transparency about the data sources that populate DSDM technologies will enable HCPs to assess important contextual factors about the data. This supports the use of data in new contexts. Transparency can be achieved by labeling the data sources on the user interface and visually representing their reliability. For example, data from the coding department could have icons alongside it, which indicate that the code is a working diagnosis or overlaps with other respiratory conditions.

Subjective Data Recorded by Patients Is Challenging for Decision-making

Despite the benefits of viewing patient-generated data, self-assessed data may be too variable for decision-making. Previous work has shown that it is challenging for patients with COPD to answer subjective questions about breathlessness and coughing [79]. To address this, patients may underreport symptoms unless there are large deviations from their baseline [79]. Unreliable reporting of symptoms impacts how data are interpreted by HCPs, which means that patients may not receive the care they require. This concern relates to the human-data interaction challenge of legibility, making data transparent and comprehensible [46].

We found that there are contexts in which subjectivity in patient-generated data is acceptable, such as clinical discussions about perceived symptoms and quality of life. Scenario 3 presented symptom data in a structured format, enabling HCPs to quickly pinpoint moments in time. Patient-generated data, in turn, becomes a useful resource for HCPs and patients to collaboratively identify personalized management strategies and goals [80]. Therefore, although quantitative symptom readings can address variability in patient reports [81], complete quantification of a patient's chronic health experience removes an important perspective. A combination of quantitative and subjective data can provide a holistic view of a patient's condition to support the development of personalized goals. However, patients may require support to understand their data

in preparation for clinic visits to maximize the value of the co-interpretation process [80].

To support patients' understanding of their data, digital technologies for self-monitoring could prompt them to input written context alongside symptom changes; for example, if symptoms deviate from a baseline. The written context prepares patients to discuss key moments in the clinic visit. However, in contexts where HCPs receive data remotely, HCPs may feel concerned if they deem themselves liable to address the content of patients' free-text notes outside of clinic visits [82]. To mitigate this, when viewing patient-generated data remotely, typed notes could be inaccessible until the HCP interacts with the patient directly. Future work is needed to explore how to connect context to symptoms without causing these concerns in HCPs.

Sporadic Data Entry by Patients Has Value

A notable challenge with patient-generated data is the perceived effort required to encourage patients to record data consistently so that health patterns can be identified [83,84]. Thus, sporadic data entry can mean that important insights are missed or rendered ambiguous [85]. For instance, sporadic data can cause challenges where complete data are required to predict exacerbations [86]. Similarly, sporadic data may indicate that a patient is too unwell to monitor their symptoms [85,87,88]. However, we found that sporadic data could have value depending on the use case.

Patients' symptomatic days were key information for HCPs, as this required some action from them. Therefore, HCPs suggested that recording data about being unwell would be a valuable insight, despite the absence of recording asymptomatic days. Patients who prefer to reduce their time thinking about their condition may also prefer to record data only when symptomatic [10]. HCPs suggested that Respire could have a filter that only displayed patients who were symptomatic and required support. However, when applying filters to data sets, HCPs may inadvertently pay less attention to patients outside the filtered subset [27]. Therefore, filters applied to the views of patient-generated data should have alerts regularly reminding the user of the applied filter.

It should be highlighted that enabling patients and HCPs to discuss health improvements is important [89,90]. Therefore, we do not argue that asymptomatic days should not be tracked as they can provide a measure, and a reminder, of how frequently patients feel *well*. Rather, we have found in contexts where there is no hard requirement to record data each day, a focus on symptomatic days alone can provide value. Future work should further explore such use cases to identify the key opportunities for sporadic data.

Support and Clear Processes Are Needed When Interacting With New Data

Concerns about responding to data can impact HCPs' desire to integrate data into their workflow [88,91-93]. This relates to the human-data interaction challenge of agency, regarding acting on data and its implications [46]. Patient-generated data present a novel opportunity to support decision-making. However, HCPs were concerned that they would be expected to instantly

investigate (lack of) data and the work involved in meeting this expectation. They felt that patients may stop acting on their health concerns as they expected HCPs to closely monitor them. In addition, HCPs had no guidance on the investigation and interpretation of patient-generated data. Aligning expectations about responding to data is important for HCPs to use data in practice [88,93]. Bardram and Frost [88] observed similar challenges raised by nurses that were responding to low mood reported by patients with bipolar disorder. This challenge highlights how wider sociotechnical considerations influence how HCPs engage with data.

Future studies should explore aligning expectations and establishing processes for responding to patient-generated data to alleviate concerns. This can be achieved by working with HCPs to understand the patient segments [94] who they wish to receive data from, the data required, and the frequency of its collection. Following this, we can collaboratively design appropriate workflows, dataflows, and digital interfaces. Our study found 4 use cases for patient-generated data: (1) supporting discussions in clinic visits, (2) monitoring acute patients to detect deterioration, (3) temporary monitoring of patients discharged from the hospital, and (4) temporary monitoring of newly diagnosed patients. Each use case may benefit from different processes, data, and visualizations. Balancing HCPs' data needs with patients' expectations in different contexts can support an understanding of how these systems can work in practice.

Limitations

This study has 2 important methodological limitations. First, the exploration of Respire consisted of the same HCPs who informed its design (except for H5 and C6). Involving the same HCPs in the design could have introduced a positive bias into the feedback, with participants potentially responding more positively to Respire [95]. The second limitation is that this research was undertaken with 2 National Health Service organizations in North West England. Their local context and ways of working have shaped our findings, which require acknowledgment when transferring the findings to other health care contexts [50].

Conclusions

By exploring data interaction scenarios with HCPs, we unearthed lessons and design implications for DSDM technologies in the context of COPD care. Although DSDM technologies can support HCPs, there are important human-data interaction and sociotechnical challenges that influence their design and deployment. These challenges are related to (1) trusting data for clinical decision-making, (2) navigating unreliable and incomplete data sets, and (3) interpreting and responding to new types of data. Further investigation of these challenges will enhance the design and deployment of effective DSDM technologies for health care. Although COPD was our area of focus, we argue that our findings have the potential to translate [50] to other areas where DSDM technologies might be used in health care.

Acknowledgments

The authors are thankful to the participants for their dedication to this study. The authors also thank Professor Mike Pearson for his support and Dr Olatunji Johnson for his advice on the creation of test data sets. This project was funded by Connected Health Cities, a Northern Health Science Alliance-led program funded by the Department of Health and delivered by a consortium of academic and National Health Service organizations across northern England. The views expressed are those of the authors.

Conflicts of Interest

None declared.

Multimedia Appendix 1

Plan for data needs workshops.

[DOCX File, 21 KB - [humanfactors_v9i2e32456_app1.docx](#)]

Multimedia Appendix 2

Plan to explore Respire.

[DOCX File, 21 KB - [humanfactors_v9i2e32456_app2.docx](#)]

Multimedia Appendix 3

Stage 1 data requirements.

[DOCX File, 20 KB - [humanfactors_v9i2e32456_app3.docx](#)]

References

1. Chronic respiratory diseases. World Health Organization. 2018. URL: https://www.who.int/health-topics/chronic-respiratory-diseases#tab=tab_1 [accessed 2020-08-01]
2. What is COPD? National Heart, Lung, and Blood Institute. URL: <https://www.nhlbi.nih.gov/health-topics/copd> [accessed 2019-06-08]

3. Mannino DM, Buist AS. Global burden of COPD: risk factors, prevalence, and future trends. *Lancet* 2007 Sep 01;370(9589):765-773. [doi: [10.1016/S0140-6736\(07\)61380-4](https://doi.org/10.1016/S0140-6736(07)61380-4)] [Medline: [17765526](https://pubmed.ncbi.nlm.nih.gov/17765526/)]
4. Hu G, Zhou Y, Tian J, Yao W, Li J, Li B, et al. Risk of COPD from exposure to biomass smoke: a metaanalysis. *Chest* 2010 Jul;138(1):20-31. [doi: [10.1378/chest.08-2114](https://doi.org/10.1378/chest.08-2114)] [Medline: [20139228](https://pubmed.ncbi.nlm.nih.gov/20139228/)]
5. Viegi G, Scognamiglio A, Baldacci S, Pistelli F, Carrozzi L. Epidemiology of chronic obstructive pulmonary disease (COPD). *Respiration* 2001;68(1):4-19. [doi: [10.1159/000050456](https://doi.org/10.1159/000050456)] [Medline: [11223724](https://pubmed.ncbi.nlm.nih.gov/11223724/)]
6. Elkington H, White P, Addington-Hall J, Higgs R, Pettinari C. The last year of life of COPD: a qualitative study of symptoms and services. *Respir Med* 2004 May;98(5):439-445 [FREE Full text] [doi: [10.1016/j.rmed.2003.11.006](https://doi.org/10.1016/j.rmed.2003.11.006)] [Medline: [15139573](https://pubmed.ncbi.nlm.nih.gov/15139573/)]
7. Miravittles M, de la Roza C, Morera J, Montemayor T, Gobartt E, Martín A, et al. Chronic respiratory symptoms, spirometry and knowledge of COPD among general population. *Respir Med* 2006 Nov;100(11):1973-1980 [FREE Full text] [doi: [10.1016/j.rmed.2006.02.024](https://doi.org/10.1016/j.rmed.2006.02.024)] [Medline: [16626950](https://pubmed.ncbi.nlm.nih.gov/16626950/)]
8. Wedzicha JA, Seemungal TA. COPD exacerbations: defining their cause and prevention. *Lancet* 2007 Sep 01;370(9589):786-796 [FREE Full text] [doi: [10.1016/S0140-6736\(07\)61382-8](https://doi.org/10.1016/S0140-6736(07)61382-8)] [Medline: [17765528](https://pubmed.ncbi.nlm.nih.gov/17765528/)]
9. Cicutto L, Brooks D, Henderson K. Self-care issues from the perspective of individuals with chronic obstructive pulmonary disease. *Patient Educ Couns* 2004 Nov;55(2):168-176. [doi: [10.1016/j.pec.2003.08.012](https://doi.org/10.1016/j.pec.2003.08.012)] [Medline: [15530751](https://pubmed.ncbi.nlm.nih.gov/15530751/)]
10. Tendedez H, Ferrario MA, McNaney R, Whittle J. Respiratory self-care: identifying current challenges and future potentials for digital technology to support people with chronic respiratory conditions. In: Proceedings of the 13th EAI International Conference on Pervasive Computing Technologies for Healthcare. 2019 Presented at: PervasiveHealth'19; May 20-23, 2019; Trento, Italy p. 129-138 URL: <https://doi.org/10.1145/3329189.3329190> [doi: [10.1145/3329189.3329190](https://doi.org/10.1145/3329189.3329190)]
11. Bauer WS, Schiffman RF. Factors influencing self-management of chronic obstructive pulmonary disease by community-dwelling adults. *West J Nurs Res* 2020 Jun;42(6):423-430. [doi: [10.1177/0193945919865532](https://doi.org/10.1177/0193945919865532)] [Medline: [31313648](https://pubmed.ncbi.nlm.nih.gov/31313648/)]
12. McManus TE, Marley AM, Baxter N, Christie SN, O'Neill HJ, Elborn JS, et al. Respiratory viral infection in exacerbations of COPD. *Respir Med* 2008 Nov;102(11):1575-1580 [FREE Full text] [doi: [10.1016/j.rmed.2008.06.006](https://doi.org/10.1016/j.rmed.2008.06.006)] [Medline: [18672353](https://pubmed.ncbi.nlm.nih.gov/18672353/)]
13. What is COPD? Asthma+Lung UK. URL: <https://www.blf.org.uk/support-for-you/copd/symptoms> [accessed 2020-10-01]
14. Nolte E, Knai C, Hofmarcher M, Conklin A, Erlner A, Elissen A, et al. Overcoming fragmentation in health care: chronic care in Austria, Germany and The Netherlands. *Health Econ Policy Law* 2012 Jan;7(1):125-146. [doi: [10.1017/S1744133111000338](https://doi.org/10.1017/S1744133111000338)] [Medline: [22221931](https://pubmed.ncbi.nlm.nih.gov/22221931/)]
15. COPD National Action Plan. National Heart, Lung, and Blood Institute. URL: https://www.nhlbi.nih.gov/sites/default/files/media/docs/COPD%20National%20Action%20Plan%20508_0.pdf [accessed 2021-12-12]
16. Soriano JB, Rodríguez-Roisin R. Chronic obstructive pulmonary disease overview: epidemiology, risk factors, and clinical presentation. *Proc Am Thorac Soc* 2011 Aug;8(4):363-367. [doi: [10.1513/pats.201102-017RM](https://doi.org/10.1513/pats.201102-017RM)] [Medline: [21816993](https://pubmed.ncbi.nlm.nih.gov/21816993/)]
17. Snell N, Strachan D, Hubbard R, Gibson J, Gruffydd-Jones K, Jarrold I. S32 Epidemiology of chronic obstructive pulmonary disease (COPD) in the UK: findings from the british lung foundation's 'respiratory health of the nation' project. *Thorax* 2016 Nov 15;71(Suppl 3):A20.1-A2A20. [doi: [10.1136/thoraxjnl-2016-209333.38](https://doi.org/10.1136/thoraxjnl-2016-209333.38)]
18. Price LC, Lowe D, Hosker HS, Anstey K, Pearson MG, Roberts CM, British Thoracic Society and the Royal College of Physicians Clinical Effectiveness Evaluation Unit (CEEu). UK National COPD Audit 2003: impact of hospital resources and organisation of care on patient outcome following admission for acute COPD exacerbation. *Thorax* 2006 Oct;61(10):837-842 [FREE Full text] [doi: [10.1136/thx.2005.049940](https://doi.org/10.1136/thx.2005.049940)] [Medline: [16449268](https://pubmed.ncbi.nlm.nih.gov/16449268/)]
19. Ebrall P. Clinical reasoning in the health professions, 3rd edition. *J Chiropr Educ* 2008;22(2):161-162.
20. Smith R. What clinical information do doctors need? *BMJ* 1996 Oct 26;313(7064):1062-1068 [FREE Full text] [doi: [10.1136/bmj.313.7064.1062](https://doi.org/10.1136/bmj.313.7064.1062)] [Medline: [8898602](https://pubmed.ncbi.nlm.nih.gov/8898602/)]
21. Dobler C. Defragmenting health care for patients with chronic obstructive pulmonary disease. *Aust Med J* 2016;9(6):185-188. [doi: [10.4066/AMJ.2016.2649](https://doi.org/10.4066/AMJ.2016.2649)]
22. Casas A, Troosters T, Garcia-Aymerich J, Roca J, Hernández C, Alonso A, members of the CHRONIC Project. Integrated care prevents hospitalisations for exacerbations in COPD patients. *Eur Respir J* 2006 Jul;28(1):123-130 [FREE Full text] [doi: [10.1183/09031936.06.00063205](https://doi.org/10.1183/09031936.06.00063205)] [Medline: [16611656](https://pubmed.ncbi.nlm.nih.gov/16611656/)]
23. Fitzpatrick G, Ellingsen G. A review of 25 years of CSCW research in healthcare: contributions, challenges and future agendas. *Comput Supported Coop Work* 2012 Jun 21;22(4-6):609-665 [FREE Full text] [doi: [10.1007/s10606-012-9168-0](https://doi.org/10.1007/s10606-012-9168-0)]
24. Tendedez H, McNaney R, Ferrario MA, Whittle J. Scoping the design space for data supported decision-making tools in respiratory care: needs, barriers and future aspirations. In: Proceedings of the 12th EAI International Conference on Pervasive Computing Technologies for Healthcare. 2018 Presented at: PervasiveHealth '18; May 21-24, 2018; New York, NY, USA p. 217-226. [doi: [10.1145/3240925.3240983](https://doi.org/10.1145/3240925.3240983)]
25. Christensen T, Grimsmo A. Instant availability of patient records, but diminished availability of patient information: a multi-method study of GP's use of electronic patient records. *BMC Med Inform Decis Mak* 2008 Mar 28;8:12 [FREE Full text] [doi: [10.1186/1472-6947-8-12](https://doi.org/10.1186/1472-6947-8-12)] [Medline: [18373858](https://pubmed.ncbi.nlm.nih.gov/18373858/)]
26. Evans RS. Electronic health records: then, now, and in the future. *Yearb Med Inform* 2018 Mar 06;25(S 01):S48-S61. [doi: [10.15265/iy-s-2016-s006](https://doi.org/10.15265/iy-s-2016-s006)]

27. Dowding D, Randell R, Gardner P, Fitzpatrick G, Dykes P, Favela J, et al. Dashboards for improving patient care: review of the literature. *Int J Med Inform* 2015 Feb;84(2):87-100. [doi: [10.1016/j.ijmedinf.2014.10.001](https://doi.org/10.1016/j.ijmedinf.2014.10.001)] [Medline: [25453274](https://pubmed.ncbi.nlm.nih.gov/25453274/)]
28. Iftikhar A, Bond R, McGilligan V, Leslie SJ, Rjoob K, Knoery C, et al. Role of dashboards in improving decision making in healthcare: review of the literature. In: *Proceedings of the 31st European Conference on Cognitive Ergonomics*. 2019 Presented at: ECCE '19; September 10-13, 2019; Belfast, UK p. 215-219 URL: <https://doi.org/10.1145/3335082.3335109> [doi: [10.1145/3335082.3335109](https://doi.org/10.1145/3335082.3335109)]
29. Litjens G, Debats O, Barentsz J, Karssemeijer N, Huisman H. Computer-aided detection of prostate cancer in MRI. *IEEE Trans Med Imaging* 2014 May;33(5):1083-1092. [doi: [10.1109/TMI.2014.2303821](https://doi.org/10.1109/TMI.2014.2303821)] [Medline: [24770913](https://pubmed.ncbi.nlm.nih.gov/24770913/)]
30. Sutton RT, Pincock D, Baumgart DC, Sadowski DC, Fedorak RN, Kroeker KI. An overview of clinical decision support systems: benefits, risks, and strategies for success. *NPJ Digit Med* 2020;3:17 [FREE Full text] [doi: [10.1038/s41746-020-0221-y](https://doi.org/10.1038/s41746-020-0221-y)] [Medline: [32047862](https://pubmed.ncbi.nlm.nih.gov/32047862/)]
31. McMenamin J, Nicholson R, Leech K. Patient Dashboard: the use of a colour-coded computerised clinical reminder in Whanganui regional general practices. *J Prim Health Care* 2011 Dec 01;3(4):307-310. [Medline: [22132384](https://pubmed.ncbi.nlm.nih.gov/22132384/)]
32. Waitman LR, Phillips IE, McCoy AB, Danciu I, Halpenny RM, Nelsen CL, et al. Adopting real-time surveillance dashboards as a component of an enterprisewide medication safety strategy. *Jt Comm J Qual Patient Saf* 2011 Jul;37(7):326-332. [doi: [10.1016/s1553-7250\(11\)37041-9](https://doi.org/10.1016/s1553-7250(11)37041-9)] [Medline: [21819031](https://pubmed.ncbi.nlm.nih.gov/21819031/)]
33. Daley K, Richardson J, James I, Chambers A, Corbett D. Clinical dashboard: use in older adult mental health wards. *Psychiatrist* 2013;37(3):85-88. [doi: [10.1192/pb.bp.111.035899](https://doi.org/10.1192/pb.bp.111.035899)]
34. Hartzler AL, IZard JP, Dalkin BL, Mikles SP, Gore JL. Design and feasibility of integrating personalized PRO dashboards into prostate cancer care. *J Am Med Inform Assoc* 2016 Jan;23(1):38-47 [FREE Full text] [doi: [10.1093/jamia/ocv101](https://doi.org/10.1093/jamia/ocv101)] [Medline: [26260247](https://pubmed.ncbi.nlm.nih.gov/26260247/)]
35. Linder JA, Schnipper JL, Tsurikova R, Yu DT, Volk LA, Melnikas AJ, et al. Electronic health record feedback to improve antibiotic prescribing for acute respiratory infections. *Am J Manag Care* 2010 Dec;16(12 Suppl HIT):e311-e319. [Medline: [21322301](https://pubmed.ncbi.nlm.nih.gov/21322301/)]
36. Dolan JG, Veazie PJ, Russ AJ. Development and initial evaluation of a treatment decision dashboard. *BMC Med Inform Decis Mak* 2013 Apr 21;13:51 [FREE Full text] [doi: [10.1186/1472-6947-13-51](https://doi.org/10.1186/1472-6947-13-51)] [Medline: [23601912](https://pubmed.ncbi.nlm.nih.gov/23601912/)]
37. Faiola A, Srinivas P, Duke J. Supporting clinical cognition: a human-centered approach to a novel ICU information visualization dashboard. *AMIA Annu Symp Proc* 2015;2015:560-569 [FREE Full text] [Medline: [26958190](https://pubmed.ncbi.nlm.nih.gov/26958190/)]
38. Kaplan B, Maxwell JA. *Qualitative Research Methods for Evaluating Computer Information Systems*. In: Anderson JG, Aydin CE, editors. *Evaluating the Organizational Impact of Healthcare Information Systems*. Health Informatics. New York, NY, USA: Springer; 2005.
39. Kaplan B. Evaluating informatics applications--clinical decision support systems literature review. *Int J Med Inform* 2001 Nov;64(1):15-37. [doi: [10.1016/s1386-5056\(01\)00183-6](https://doi.org/10.1016/s1386-5056(01)00183-6)] [Medline: [11673100](https://pubmed.ncbi.nlm.nih.gov/11673100/)]
40. Kaplan B. Evaluating informatics applications--some alternative approaches: theory, social interactionism, and call for methodological pluralism. *Int J Med Inform* 2001 Nov;64(1):39-56. [doi: [10.1016/s1386-5056\(01\)00184-8](https://doi.org/10.1016/s1386-5056(01)00184-8)] [Medline: [11673101](https://pubmed.ncbi.nlm.nih.gov/11673101/)]
41. Vogwill V, Reeves S. Challenges of information exchange between nurses and physicians in multidisciplinary team meetings. *J Interprof Care* 2008 Dec;22(6):664-667. [doi: [10.1080/13561820802114772](https://doi.org/10.1080/13561820802114772)] [Medline: [19012148](https://pubmed.ncbi.nlm.nih.gov/19012148/)]
42. Liu J, Wyatt JC, Altman DG. Decision tools in health care: focus on the problem, not the solution. *BMC Med Inform Decis Mak* 2006 Jan 20;6:4 [FREE Full text] [doi: [10.1186/1472-6947-6-4](https://doi.org/10.1186/1472-6947-6-4)] [Medline: [16426446](https://pubmed.ncbi.nlm.nih.gov/16426446/)]
43. Favela J, Tentori M, Gonzalez VM. Ecological validity and pervasiveness in the evaluation of ubiquitous computing technologies for health care. *Int J Hum Comput Interact* 2010 Apr 27;26(5):414-444. [doi: [10.1080/10447311003719896](https://doi.org/10.1080/10447311003719896)]
44. Ackerman M, Prilla M, Sary C, Herrmann TA, Goggins S. *Designing Healthcare That Works: A Sociotechnical Approach*. Cambridge, MA, USA: Academic Press; Nov 13, 2017.
45. Victorelli EZ, Dos Reis JC, Hornung H, Prado AB. Understanding human-data interaction: literature review and recommendations for design. *Int J Hum Comput Stud* 2020 Feb;134:13-32. [doi: [10.1016/j.ijhcs.2019.09.004](https://doi.org/10.1016/j.ijhcs.2019.09.004)]
46. Mortier R, Haddadi H, Henderson T, McAuley D, Crowcroft J. Human-data interaction: the human face of the data-driven society. *SSRN Journal* 2014 Oct. [doi: [10.2139/ssrn.2508051](https://doi.org/10.2139/ssrn.2508051)]
47. Sarcevic A, Marsic I, Burd RS. Chapter 7 - Dashboard Design for Improved Team Situation Awareness in Time-Critical Medical Work: Challenges and Lessons Learned. In: Ackerman MS, Goggins SP, Prilla M, Sary C, editors. *Designing Healthcare That Works: A Sociotechnical Approach*. Cambridge, MA, USA: Academic Press; 2018:113-131.
48. Bardram JE, Nørskov N. A context-aware patient safety system for the operating room. In: *Proceedings of the 10th International Conference on Ubiquitous Computing*. 2008 Presented at: UbiComp '08; September 21-24, 2008; Seoul, Korea p. 272-281. [doi: [10.1145/1409635.1409672](https://doi.org/10.1145/1409635.1409672)]
49. Sarcevic A, Weibel N, Hollan JD, Burd RS. A paper-digital interface for information capture and display in time-critical medical work. In: *Proceedings of the 6th International Conference on Pervasive Computing Technologies for Healthcare (PervasiveHealth) and Workshops*. 2012 Presented at: PervasiveHealth '12; May 21-24, 2012; San Diego, CA, USA p. 17-24. [doi: [10.4108/icst.pervasivehealth.2012.248674](https://doi.org/10.4108/icst.pervasivehealth.2012.248674)]

50. Hayes GR. The relationship of action research to human-computer interaction. *ACM Trans Comput Hum Interact* 2011 Jul 01;18(3):1-20. [doi: [10.1145/1993060.1993065](https://doi.org/10.1145/1993060.1993065)]
51. Bardram JE. Scenario-based design of cooperative systems re-designing a hospital information system in Denmark. *Group Decis Negot* 2000;9(3):237-250. [doi: [10.1023/A:1008743509338](https://doi.org/10.1023/A:1008743509338)]
52. Dahl Y, Alsos OA, Svanæs D. Fidelity considerations for simulation-based usability assessments of mobile ICT for hospitals. *Int J Hum Comput Interact* 2010 Apr 27;26(5):445-476. [doi: [10.1080/10447311003719938](https://doi.org/10.1080/10447311003719938)]
53. Carroll JM. Scenario-based design: envisioning work and technology in system development. New York, NY, USA: John Wiley & Sons; 1995.
54. Sanders EB, Stappers PJ. Co-creation and the new landscapes of design. *CoDesign* 2008 Mar;4(1):5-18. [doi: [10.1080/15710880701875068](https://doi.org/10.1080/15710880701875068)]
55. Corhay JL, Dang DN, Van Cauwenberge H, Louis R. Pulmonary rehabilitation and COPD: providing patients a good environment for optimizing therapy. *Int J Chron Obstruct Pulmon Dis* 2014;9:27-39 [FREE Full text] [doi: [10.2147/COPD.S52012](https://doi.org/10.2147/COPD.S52012)] [Medline: [24368884](https://pubmed.ncbi.nlm.nih.gov/24368884/)]
56. Escarrabill J. Discharge planning and home care for end-stage COPD patients. *Eur Respir J* 2009 Aug;34(2):507-512 [FREE Full text] [doi: [10.1183/09031936.00146308](https://doi.org/10.1183/09031936.00146308)] [Medline: [19648525](https://pubmed.ncbi.nlm.nih.gov/19648525/)]
57. Barreiro TJ, Perillo I. An approach to interpreting spirometry. *Am Fam Physician* 2004 Mar 01;69(5):1107-1114 [FREE Full text] [Medline: [15023009](https://pubmed.ncbi.nlm.nih.gov/15023009/)]
58. Johnson TP. Snowball Sampling: Introduction. In: Wiley StatsRef: Statistics Reference Online. Hoboken, NJ, USA: John Wiley & Sons; Sep 29, 2014.
59. Rehman AA, Alharthi K. An introduction to research paradigms. *Int J Educ Investig* 2016 Oct;3(8):51-59. [doi: [10.4018/978-1-5225-1738-2.ch001](https://doi.org/10.4018/978-1-5225-1738-2.ch001)]
60. Vaismoradi M, Turunen H, Bondas T. Content analysis and thematic analysis: implications for conducting a qualitative descriptive study. *Nurs Health Sci* 2013 Sep;15(3):398-405. [doi: [10.1111/nhs.12048](https://doi.org/10.1111/nhs.12048)] [Medline: [23480423](https://pubmed.ncbi.nlm.nih.gov/23480423/)]
61. Braun V, Clarke V. Using thematic analysis in psychology. *Qual Res Psychol* 2006 Jan;3(2):77-101. [doi: [10.1191/1478088706qp0630a](https://doi.org/10.1191/1478088706qp0630a)]
62. Wharton C, Riemann J, Lewis C, Poison P. The cognitive walkthrough method: a practitioner's guide. In: Nielsen J, Mack RL, editors. Usability inspection methods. New York, NY, USA: John Wiley & Sons; Jun 1994:105-140.
63. International Statistical Classification of Diseases and Related Health Problems (ICD). World Health Organization. URL: <https://www.who.int/classifications/icd/en/> [accessed 2021-12-12]
64. Maurer J, Rebbapragada V, Borson S, Goldstein R, Kunik ME, Yohannes AM, ACCP Workshop Panel on Anxiety and Depression in COPD. Anxiety and depression in COPD: current understanding, unanswered questions, and research needs. *Chest* 2008 Oct;134(4 Suppl):43S-56S [FREE Full text] [doi: [10.1378/chest.08-0342](https://doi.org/10.1378/chest.08-0342)] [Medline: [18842932](https://pubmed.ncbi.nlm.nih.gov/18842932/)]
65. Heinzer MM, Bish C, Detwiler R. Acute dyspnea as perceived by patients with chronic obstructive pulmonary disease. *Clin Nurs Res* 2003 Feb;12(1):85-101. [doi: [10.1177/1054773803238742](https://doi.org/10.1177/1054773803238742)] [Medline: [12583501](https://pubmed.ncbi.nlm.nih.gov/12583501/)]
66. Xu W, Collet J, Shapiro S, Lin Y, Yang T, Platt RW, et al. Independent effect of depression and anxiety on chronic obstructive pulmonary disease exacerbations and hospitalizations. *Am J Respir Crit Care Med* 2008 Nov 01;178(9):913-920. [doi: [10.1164/rccm.200804-619OC](https://doi.org/10.1164/rccm.200804-619OC)] [Medline: [18755925](https://pubmed.ncbi.nlm.nih.gov/18755925/)]
67. Rezaee M, Ward CE, Nunez B, Rezaee DA, Dittkoff J, Halalau A. Examining 30-day COPD readmissions through the emergency department. *Int J Chron Obstruct Pulmon Dis* 2018;13:109-120 [FREE Full text] [doi: [10.2147/COPD.S147796](https://doi.org/10.2147/COPD.S147796)] [Medline: [29343950](https://pubmed.ncbi.nlm.nih.gov/29343950/)]
68. Cicourel AV. The integration of distributed knowledge in collaborative medical diagnosis. In: Galegher J, Kraut RE, Egido C, editors. Intellectual Teamwork: Social and Technological Foundations of Cooperative Work. Mahwah, NJ, USA: Lawrence Erlbaum Associates; 1990.
69. Ellingsen G. The role of trust in knowledge management: a case study of physicians at work at the University Hospital of Northern Norway. *Inform Sci* 2003;6:193-207. [doi: [10.28945/526](https://doi.org/10.28945/526)]
70. Ellingsen G. Knowledge Work in Hospitals. In: Proceedings of the IFIP TC8/WG8.2 Working Conference on Global and Organizational Discourse about Information Technology. 2002 Presented at: IFIP TC8/WG8.2 '02; December 12-14, 2002; Barcelona, Spain p. 465-484. [doi: [10.1007/978-0-387-35634-1_22](https://doi.org/10.1007/978-0-387-35634-1_22)]
71. Jirotko M, Procter R, Hartswood M, Slack R, Simpson A, Coopmans C, et al. Collaboration and trust in healthcare innovation: the eDiaMoND case study. *Comput Supported Coop Work* 2005 Sep 14;14(4):369-398 [FREE Full text] [doi: [10.1007/s10606-005-9001-0](https://doi.org/10.1007/s10606-005-9001-0)]
72. Anderson A, Huttenlocher D, Kleinberg J, Leskovec J. Steering user behavior with badges. In: Proceedings of the 22nd international conference on World Wide Web. 2013 Presented at: WWW '13; May 13-17, 2013; Rio de Janeiro, Brazil p. 95-106 URL: <https://doi.org/10.1145/2488388.2488398> [doi: [10.1145/2488388.2488398](https://doi.org/10.1145/2488388.2488398)]
73. Knowles B. Emerging trust implications of data-rich systems. *IEEE Pervasive Comput* 2016 Oct 25;15(4):76-84. [doi: [10.1109/mprv.2016.68](https://doi.org/10.1109/mprv.2016.68)]
74. Axelrod L, Fitzpatrick G, Henwood F, Cassell J, Smith H, Nicholson A, et al. Data recording in primary care field studies: patient records enhancement project. In: Proceedings of the 5th International Conference on Pervasive Computing

- Technologies for Healthcare (PervasiveHealth) and Workshops. 2011 Presented at: PervasiveHealth '11; May 23-26, 2011; Dublin, Ireland p. 563-567. [doi: [10.4108/icst.pervasivehealth.2011.246073](https://doi.org/10.4108/icst.pervasivehealth.2011.246073)]
75. Sunderland R. Inaccurate coding corrupts medical information. *Arch Dis Child* 1985 Jun;60(6):593-594 [FREE Full text] [doi: [10.1136/adc.60.6.593](https://doi.org/10.1136/adc.60.6.593)] [Medline: [4015180](https://pubmed.ncbi.nlm.nih.gov/4015180/)]
76. Burns EM, Rigby E, Mamidanna R, Bottle A, Aylin P, Ziprin P, et al. Systematic review of discharge coding accuracy. *J Public Health (Oxf)* 2012 Mar;34(1):138-148 [FREE Full text] [doi: [10.1093/pubmed/fdr054](https://doi.org/10.1093/pubmed/fdr054)] [Medline: [21795302](https://pubmed.ncbi.nlm.nih.gov/21795302/)]
77. Berg M, Goorman E. The contextual nature of medical information. *Int J Med Inform* 1999 Dec;56(1-3):51-60. [doi: [10.1016/s1386-5056\(99\)00041-6](https://doi.org/10.1016/s1386-5056(99)00041-6)] [Medline: [10659934](https://pubmed.ncbi.nlm.nih.gov/10659934/)]
78. Berg M, Langenberg C, vd Berg I, Kwakkernaat J. Considerations for sociotechnical design: experiences with an electronic patient record in a clinical context. *Int J Med Inform* 1998;52(1-3):243-251. [doi: [10.1016/s1386-5056\(98\)00143-9](https://doi.org/10.1016/s1386-5056(98)00143-9)] [Medline: [9848421](https://pubmed.ncbi.nlm.nih.gov/9848421/)]
79. Nadarajah SG, Pedersen PW, Hougaard BI, Knoche H. Am I coughing more than usual? Patient reflections and user needs on tracking COPD data in a telehealth system. In: Proceedings of the 4th International Workshop on Multimedia for Personal Health & Health Care. 2019 Presented at: HealthMedia '19; October 21, 2019; Nice, France p. 2-8 URL: <https://doi.org/10.1145/3347444.3356237> [doi: [10.1145/3347444.3356237](https://doi.org/10.1145/3347444.3356237)]
80. Mentis HM, Komlodi A, Schrader K, Phipps M, Gruber-Baldini A, Yarbrough K, et al. Crafting a view of self-tracking data in the clinical visit. In: Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems. 2017 Presented at: CHI '17; May 6-11, 2017; Denver, CO, USA p. 5800-5812. [doi: [10.1145/3025453.3025589](https://doi.org/10.1145/3025453.3025589)]
81. Velardo C, Shah SA, Gibson O, Clifford G, Heneghan C, Rutter H, EDGE COPD Team. Digital health system for personalised COPD long-term management. *BMC Med Inform Decis Mak* 2017 Feb 20;17(1):19 [FREE Full text] [doi: [10.1186/s12911-017-0414-8](https://doi.org/10.1186/s12911-017-0414-8)] [Medline: [28219430](https://pubmed.ncbi.nlm.nih.gov/28219430/)]
82. Marcu G, Bardram JE, Gabrielli S. A framework for overcoming challenges in designing persuasive monitoring and feedback systems for mental illness. In: Proceedings of the 5th International Conference on Pervasive Computing Technologies for Healthcare (PervasiveHealth) and Workshops. 2011 Presented at: PervasiveHealth '11; May 23-26, 2011; Dublin, Ireland p. 1-8. [doi: [10.4108/icst.pervasivehealth.2011.246097](https://doi.org/10.4108/icst.pervasivehealth.2011.246097)]
83. Stone AA, Shiffman S, Schwartz JE, Broderick JE, Hufford MR. Patient compliance with paper and electronic diaries. *Control Clin Trials* 2003 Apr;24(2):182-199. [doi: [10.1016/s0197-2456\(02\)00320-3](https://doi.org/10.1016/s0197-2456(02)00320-3)] [Medline: [12689739](https://pubmed.ncbi.nlm.nih.gov/12689739/)]
84. Stone AA, Shiffman S, Schwartz JE, Broderick JE, Hufford MR. Patient non-compliance with paper diaries. *BMJ* 2002 May 18;324(7347):1193-1194 [FREE Full text] [doi: [10.1136/bmj.324.7347.1193](https://doi.org/10.1136/bmj.324.7347.1193)] [Medline: [12016186](https://pubmed.ncbi.nlm.nih.gov/12016186/)]
85. West P, Van Kleek M, Giordano R, Weal MJ, Shadbolt N. Common barriers to the use of patient-generated data across clinical settings. In: Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems. 2018 Presented at: CHI '18; April 29-May 5, 2022; New Orleans, LA, USA p. 1-13 URL: <https://doi.org/10.1145/3173574.3174058> [doi: [10.1145/3173574.3174058](https://doi.org/10.1145/3173574.3174058)]
86. Anastasiou A, Kocsis O, Moustakas K. Exploring machine learning for monitoring and predicting severe asthma exacerbations. In: Proceedings of the 10th Hellenic Conference on Artificial Intelligence. 2018 Presented at: SETN '18; July 9-12, 2018; Patras, Greece p. 1-6 URL: <https://doi.org/10.1145/3200947.3201036> [doi: [10.1145/3200947.3201036](https://doi.org/10.1145/3200947.3201036)]
87. West P, Giordano R, Van Kleek M, Shadbolt N. Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems. 2016 Presented at: CHI '16; May 7-12, 2016; San Jose, CA, USA p. 3066-3078 URL: <https://doi.org/10.1145/2858036.2858445> [doi: [10.1145/2858036.2858445](https://doi.org/10.1145/2858036.2858445)]
88. Bardram JE, Frost MM. Chapter 10 - Double-Loop Health Technology Enabling Socio-technical Design of Personal Health Technology in Clinical Practice. In: Ackerman MS, Goggins SP, Herrmann T, Prilla M, Sary C, editors. *Designing Healthcare That Works: A Sociotechnical Approach*. Cambridge, MA, USA: Academic Press; 2018:167-186.
89. Akinsiku A, Avellino I, Graham Y, Mentis HM. It's not just the movement: experiential information needed for stroke telerehabilitation. In: Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems. 2021 Presented at: CHI '21; May 8-13, 2021; Yokohama, Japan p. 1-12 URL: <https://doi.org/10.1145/3411764.3445663> [doi: [10.1145/3411764.3445663](https://doi.org/10.1145/3411764.3445663)]
90. Mentis HM, Shewbridge R, Powell S, Armstrong M, Fishman P, Shulman L. Co-interpreting movement with sensors: assessing Parkinson's patients' deep brain stimulation programming. *Human Comput Interact* 2015 Aug 25;31(3-4):227-260. [doi: [10.1080/07370024.2015.1073592](https://doi.org/10.1080/07370024.2015.1073592)]
91. Cai CJ, Reif E, Hegde N, Hipp J, Kim B, Smilkov D, et al. Human-centered tools for coping with imperfect algorithms during medical decision-making. In: Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems. 2019 Presented at: CHI '19; May 4-9, 2019; Glasgow, UK p. 1-14 URL: <https://doi.org/10.1145/3290605.3300234> [doi: [10.1145/3290605.3300234](https://doi.org/10.1145/3290605.3300234)]
92. Tendedez H, Ferrario MA, McNaney R. 'The issue with that sort of data?': clinicians' accountability concerns around COPD self-monitoring tools. In: Conference Companion Publication of the 2019 on Computer Supported Cooperative Work and Social Computing. 2019 Presented at: CSCW '19; November 9-13, 2019; Austin, TX, USA p. 382-386 URL: <https://doi.org/10.1145/3311957.3359464> [doi: [10.1145/3311957.3359464](https://doi.org/10.1145/3311957.3359464)]
93. Ye J. The impact of electronic health record-integrated patient-generated health data on clinician burnout. *J Am Med Inform Assoc* 2021 Apr 23;28(5):1051-1056 [FREE Full text] [doi: [10.1093/jamia/ocab017](https://doi.org/10.1093/jamia/ocab017)] [Medline: [33822095](https://pubmed.ncbi.nlm.nih.gov/33822095/)]

94. Vuik SI, Mayer EK, Darzi A. Patient segmentation analysis offers significant benefits for integrated care and support. *Health Aff (Millwood)* 2016 May 01;35(5):769-775. [doi: [10.1377/hlthaff.2015.1311](https://doi.org/10.1377/hlthaff.2015.1311)] [Medline: [27140981](https://pubmed.ncbi.nlm.nih.gov/27140981/)]
95. Norton M, Mochon D, Ariely D. The IKEA effect: when labor leads to love. *J Consum Psychol* 2012 Jul;22(3):453-460 [FREE Full text] [doi: [10.1016/j.jcps.2011.08.002](https://doi.org/10.1016/j.jcps.2011.08.002)]

Abbreviations

COPD: chronic obstructive pulmonary disease

DSDM: data-supported decision-making

GP: general practitioner

HCP: health care professional

REST: Representational State Transfer

Edited by A Kushniruk; submitted 28.07.21; peer-reviewed by S Bhattacharjee, A Lau, A Joseph; comments to author 07.11.21; revised version received 07.02.22; accepted 21.03.22; published 06.05.22.

Please cite as:

Tendedez H, Ferrario MA, McNaney R, Gradinar A

Exploring Human-Data Interaction in Clinical Decision-making Using Scenarios: Co-design Study

JMIR Hum Factors 2022;9(2):e32456

URL: <https://humanfactors.jmir.org/2022/2/e32456>

doi: [10.2196/32456](https://doi.org/10.2196/32456)

PMID: [35522463](https://pubmed.ncbi.nlm.nih.gov/35522463/)

©Helena Tendedez, Maria-Angela Ferrario, Roisin McNaney, Adrian Gradinar. Originally published in *JMIR Human Factors* (<https://humanfactors.jmir.org>), 06.05.2022. This is an open-access article distributed under the terms of the Creative Commons Attribution License (<https://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in *JMIR Human Factors*, is properly cited. The complete bibliographic information, a link to the original publication on <https://humanfactors.jmir.org>, as well as this copyright and license information must be included.

Original Paper

Combining Persuasive System Design Principles and Behavior Change Techniques in Digital Interventions Supporting Long-term Weight Loss Maintenance: Design and Development of eCHANGE

Rikke Aune Asbjørnsen^{1,2,3}, MSc; Jøran Hjelmesæth^{4,5}, PhD; Mirjam Lien Smedsrød⁶, MSc; Jobke Wentzel^{1,7}, PhD; Marianne Ollivier³, MSc; Matthew M Clark⁸, PhD; Julia E W C van Gemert-Pijnen^{1,9*}, PhD; Lise Solberg Nes^{3,8,10*}, PhD

¹Center for eHealth and Wellbeing Research, Department of Psychology, Health and Technology, University of Twente, Enschede, Netherlands

²Research and Innovation Department, Vestfold Hospital Trust, Tønsberg, Norway

³Department of Digital Health Research, Division of Medicine, Oslo University Hospital, Oslo, Norway

⁴Morbid Obesity Center, Vestfold Hospital Trust, Tønsberg, Norway

⁵Department of Endocrinology, Morbid Obesity and Preventive Medicine, Institute of Clinical Medicine, University of Oslo, Oslo, Norway

⁶Collaborative Care Unit, Sørlandet Hospital Trust, Kristiansand, Norway

⁷Research Group IT Innovations in Health Care, Windesheim University of Applied Sciences, Zwolle, Netherlands

⁸Department of Psychiatry & Psychology, College of Medicine & Science, Mayo Clinic, Rochester, MN, United States

⁹University of Waterloo, Waterloo, ON, Canada

¹⁰Institute of Clinical Medicine, Faculty of Medicine, University of Oslo, Oslo, Norway

* these authors contributed equally

Corresponding Author:

Rikke Aune Asbjørnsen, MSc

Center for eHealth and Wellbeing Research

Department of Psychology, Health and Technology

University of Twente

De Zulf 10

Enschede, 7522 NJ

Netherlands

Phone: 31 534899111

Email: r.a.asbjornsen@utwente.nl

Abstract

Background: Long-term weight maintenance after weight loss is challenging, and innovative solutions are required. Digital technologies can support behavior change and, therefore, have the potential to be an effective tool for weight loss maintenance. However, to create meaningful and effective digital behavior change interventions that support end user values and needs, a combination of persuasive system design (PSD) principles and behavior change techniques (BCTs) might be needed.

Objective: This study aimed to investigate how an evidence-informed digital behavior change intervention can be designed and developed by combining PSD principles and BCTs into design features to support end user values and needs for long-term weight loss maintenance.

Methods: This study presents a concept for how PSD principles and BCTs can be translated into design features by combining design thinking and Agile methods to develop and deliver an evidence-informed digital behavior change intervention aimed at supporting weight maintenance. Overall, 45 stakeholders participated in the systematic and iterative development process comprising co-design workshops, prototyping, Agile development, and usability testing. This included prospective end users (n=17, 38%; ie, people with obesity who had lost $\geq 8\%$ of their weight), health care providers (n=9, 20%), healthy volunteers (n=4, 9%), a service designer (n=1, 2%), and stakeholders from the multidisciplinary research and development team (n=14, 31%; ie, software developers; digital designers; and eHealth, behavior change, and obesity experts). Stakeholder input on how to operationalize the design features and optimize the technology was examined through formative evaluation and qualitative analyses using rapid and in-depth analysis approaches.

Results: A total of 17 design features combining PSD principles and BCTs were identified as important to support end user values and needs based on stakeholder input during the design and development of eCHANGE, a digital intervention to support long-term weight loss maintenance. The design features were combined into 4 main intervention components: *Week Plan, My Overview, Knowledge and Skills*, and *Virtual Coach and Smart Feedback System*. To support a healthy lifestyle and continued behavior change to maintain weight, PSD principles such as *tailoring, personalization, self-monitoring, reminders, rewards, rehearsal, praise, and suggestions* were combined and implemented into the design features together with BCTs from the clusters of *goals and planning, feedback and monitoring, social support, repetition and substitution, shaping knowledge, natural consequences, associations, antecedents, identity, and self-belief*.

Conclusions: Combining and implementing PSD principles and BCTs in digital interventions aimed at supporting sustainable behavior change may contribute to the design of engaging and motivating interventions in line with end user values and needs. As such, the design and development of the eCHANGE intervention can provide valuable input for future design and tailoring of evidence-informed digital interventions, even beyond digital interventions in support of health behavior change and long-term weight loss maintenance.

Trial Registration: ClinicalTrials.gov NCT04537988; <https://clinicaltrials.gov/ct2/show/NCT04537988>

(*JMIR Hum Factors* 2022;9(2):e37372) doi:[10.2196/37372](https://doi.org/10.2196/37372)

KEYWORDS

eHealth; weight loss maintenance; behavior change; persuasive technology; digital health interventions; design thinking; co-design; Agile development; human-centered design; mobile phone

Introduction

Background

Healthy lifestyle and behavior changes are difficult to initiate but even more challenging to sustain over time [1]. There has been an increase in lifestyle-related diseases such as obesity, cardiovascular diseases, and diabetes worldwide [2-4]. The estimated number of people with obesity nearly tripled between 1975 and 2016, with >650 million adults with obesity (BMI \geq 30 kg/m²) worldwide in 2016 [5]. Similarly, the number of people with cardiovascular diseases increased by >50% in the period from 1990 to 2019 [6], and people living with diabetes increased by >60% in the period from 2009 to 2019 [7].

For people with obesity, weight loss may be difficult; however, maintaining weight after weight loss appears to be even more of a challenge [8], and several factors (eg, environmental, biological, behavioral, and cognitive) [9-12] contribute to the complexity of this health problem. The impact and burden of obesity on an individual are substantial and multifaceted, as it can affect both health and well-being, with an increased risk of medical conditions, premature mortality, and reduced quality of life [13,14].

The Challenge of Weight Loss Maintenance

Lifestyle interventions focusing on diet and physical activity, as well as behavioral and cognitive strategies, are widely recommended for obesity management [12,15-20]. However, even when weight loss is achieved, mechanisms such as increased hunger, reduced energy expenditure, and reduced satiety frequently contribute weight regain [10,17]. As much as 30% to 50% of the initial weight that is lost during lifestyle interventions is often regained during the subsequent 2 to 3 years [21], and few manage to maintain their lost weight in the long term [22]. With several factors contributing to weight regain [9,10,23], solving the weight loss maintenance challenge appears to be a complicated endeavor.

Health behaviors and self-regulation play central roles in weight loss and weight loss maintenance [24,25]. Implementing sustainable and feasible behavior change and self-regulation strategies into daily life takes time and effort, and finding ways of initiating and maintaining behavior change over a long period is a complex undertaking [26].

Health behaviors also change over time and are often person and context related (eg, individual motives, habits, and social and environmental factors) [23,26]. With many people failing to maintain weight after initial weight loss, innovative approaches for long-term behavior change and weight loss maintenance are needed [10,27].

Advantages and Challenges of Digital Interventions

Digital interventions are increasingly being used to promote healthy lifestyles and improve health outcomes [28-31] and can be an accessible and feasible way of supporting behavior change through its availability and scalability. A digital health (ie, eHealth) intervention can be defined as a digital technology focusing on intervening in an existing situation aiming to change health behavior [29]. The potential of evidence-informed digital technologies supporting weight loss maintenance may be significant, as digital interventions can overcome the *time and place barrier* and adapt to a person's context, needs, and preferences [32], consequently offering people the support needed for sustained behavior change [29,33].

Research exploring technologies for weight loss maintenance support is still at an early stage, and little is known about their long-term effectiveness, as few evidence-based eHealth interventions are available [34,35]. In addition, little is known about potential changes in user needs over time and how digital technologies can support end users in maintaining weight in an optimal manner for long-term behavior change [32]. In fact, knowledge is lacking on how to translate persuasive system design (PSD) principles and behavior change techniques (BCTs) into design features when developing digital interventions aimed

at facilitating continued health behavior change and weight loss maintenance [35].

Development of Innovative Solutions for Sustainable Behavior Change

An important factor for the success of digital health research, development, and implementation is the early involvement of end users and other key stakeholders in the design and formative evaluation of a product or technology [29,36-39]. Human-centered design approaches such as design thinking [40,41] can be combined with principles from Agile software development [42], stimulating the *collective creativity* of end users and other key stakeholders (eg, designers, developers, researchers, and experts) for the rapid development and evaluation of digital health interventions [31,43-45].

Design thinking is an approach that emphasizes understanding and empathy with end users, multidisciplinary collaboration, and iterative involvement of stakeholders through generation of creative ideas and action-oriented rapid prototyping to create desirable, feasible, and viable innovative solutions [40,46]. Agile software development is a flexible approach that emphasizes active stakeholder involvement through rapid iterations to test assumptions and validate possible solutions to quickly learn and adapt to changes in needs [42]. Combining these design and development methods could therefore be a time- and cost-effective approach to explore and validate whether an innovative solution is desirable (ie, what users and stakeholders want) [40,42,43,47] and whether it solves the right problem (ie, meet user and stakeholder needs) for a *problem-solution fit* [48].

Translating PSD Principles and BCTs Into Design Features

To support long-lasting behavior change through digital technologies, the design needs to support the user in adopting sustainable behaviors related to their individual goals and values [31,32]. Although goals relate to something a person would like to achieve, values can refer to *what a person considers important in life* [49], reflecting end users' ideals or interests [50]. As such, values can be defined as the *main drivers of behaviors or high-level needs* [32]. Therefore, finding a *solution* that can help address end users' goals by taking their key values and needs into account is important to create meaningful and effective eHealth technologies in support of continued health behavior change.

Complex interventions, such as digital behavior change interventions, usually consist of many active ingredients or interactive components [51-54] and can be designed to facilitate motivation and adherence to healthy behaviors [35]. PSD principles and BCTs can be such *active ingredients* or *building blocks* of digital behavior change interventions. To design motivating and effective digital technologies, integrated PSD principles and BCTs should match end user values and needs [32]. PSD principles are designed to influence users' attitudes and behaviors [55] and can be applied to match user profiles to motivate and trigger health behavior change in the design of technologies [35,55,56]. By contrast, BCTs are designed to alter or redirect causal processes that regulate behavior [51,52], can

be applied to any intervention focusing on behavior change to improve the health and well-being of people [29], and can be applied to daily life without technology [32]. As such, the PSD principles and BCTs overlap and complement each other [29,35]. Theoretical principles from the PSD model by Oinas-Kukkonen [55] and BCTs from the Behavior Change Taxonomy by Michie et al [52] can be translated into design features during the design and development of digital behavior change interventions to meet end user values and needs [29,32]. Such combined PSD [55] and BCT [52] features can be embedded in a digital application *with the specific aim of forming, altering, or reinforcing healthy attitudes and behaviors* [57].

However, research involving digital behavior change interventions often fails to clearly show how theories and techniques of behavior change have been combined and applied to design and practical delivery forms [31,35,57-61]. There appears to be a lack of theoretical frameworks and specifications of design features when reporting on digital interventions. It is often unclear which design features, PSD principles, and behavior change strategies are most effective in meeting end user values and needs and how they influence health-related outcomes, including weight loss maintenance [29,35,52,62-64].

In response to these gaps, the current research group performed a scoping review and a qualitative study aimed at identifying PSD principles and BCTs from eHealth interventions applied in existing weight loss maintenance research [35], as well as key values and needs related to *what* people want for maintaining weight and *why* [32]. The scoping review [35] identified PSD principles [55] (eg, *self-monitoring, reminders, rewards, tailoring, personalization, and praise*) and BCT clusters [52] (eg, *feedback and monitoring, goals and planning, repetition and substitutions, social support, associations, and shaping knowledge*) applied in eHealth interventions to stimulate adherence, motivation, and weight loss maintenance. The technology characteristics of existing eHealth weight loss maintenance interventions were usually supported by mobile phone technology, sometimes in combination with an activity tracker and wireless scale [35]. The subsequent qualitative study (ie, individual and focus group interviews) identified key end user values of people with the aim of maintaining weight loss in the long term (ie, *autonomy, self-management, motivation, personalized care, happiness, health, feel supported, and positive self-image*), as well as PSD principles and BCTs that might be essential to include in eHealth interventions to meet end user values and needs [32]. The findings indicated that the most successful and promising eHealth weight loss maintenance interventions entailed a combination of both PSD principles and BCTs [32,35]. The studies overlapped in findings [32,35]; however, some less frequently applied PSD principles (eg, *rehearsal*) and BCTs (eg, *identity, self-belief, and natural consequences*) were identified in the qualitative study, which might be of importance to support end user values and needs to prevent weight regain in the long term [32]. The qualitative study also highlighted the *tailoring* and *personalization* of digital interventions to address the often multifaceted and dynamic changes in individual needs over time (eg, related to behaviors,

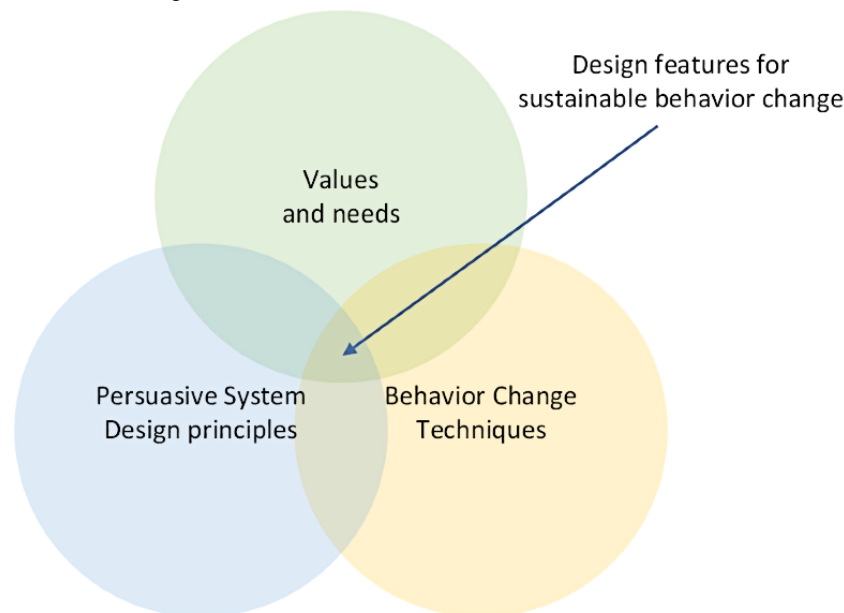
thoughts, and emotions) for continued health behavior change [32].

The results from these recent studies strongly suggest that the process of translating end user values into *design features*, as well as the exploration of how PSD principles and BCTs can be combined and implemented, are important parts of the design and development processes when aiming to create motivating and engaging digital interventions to support sustained behavior change and weight loss maintenance.

Objectives

The overall aim of this study was to investigate *how* digital technology can meet end users' values and needs by exploring and validating design features (ie, the combination and implementation of PSD principles and BCTs) through iterative design, development, and formative evaluation of a digital intervention called eCHANGE.

Figure 1. Design features for sustainable behavior change.



Methods

The Double Diamond Framework and the Center for eHealth Research and Disease Management Roadmap

To develop and deliver a digital intervention aimed at supporting long-term weight maintenance following weight loss (ie, eCHANGE; ClinicalTrials.gov NCT04537988), the process was guided by the Double Diamond (ie, design thinking process) [65,66] and the Center for eHealth Research and Disease Management (CeHRes) roadmap [29,36] (Figure 2 and Multimedia Appendix 1 [36,66]). In addition, refer to the study by Asbjørnsen et al [32] for additional background and details. To improve the uptake and impact of digital technology [29], multidisciplinary collaboration, human-centered design approaches such as design thinking and service design [41,46], Agile software development principles [42], and formative evaluation were integrated into a systematic and iterative

The following describes how 2 theoretical frameworks [52,55] targeting motivation and behavior change can be combined with applied innovation methodologies, such as design thinking [41,65] and Agile development [42], during the design of digital behavior change interventions. The findings show the development of digital technology aimed at facilitating sustainable behavior change to maintain weight after weight loss. The main research question of this study was as follows: how can an evidence-informed digital behavior change intervention, combining and implementing PSD principles and BCTs into design features, be designed and developed to support end user values and needs for long-term weight loss maintenance?

For this study, design features for (sustainable) behavior change were defined as the combination of *PSD* principles and *BCTs* embedded in a digital intervention, with the specific aim of supporting end users' (ie, target users) values and needs to facilitate sustainable health behavior change (Figure 1).

development process to optimize the fit between values and needs, technology, and context [36].

This study focused mainly on the second diamond of the Double Diamond process (ie, *develop* and *deliver*; Figure 2) [66], particularly the *Design* and *Operationalization* phases of the CeHRes Roadmap, including *formative evaluation* to gather input to improve the intervention [29,36].

During the *Design* phase [36], the aim was to explore how digital technology can be *developed* and to validate which design features meet end users' values and needs to support long-term weight loss maintenance. Knowledge, insights, and ideas from previous studies [32,35] were translated into prototypes through co-design workshops [41], rapid prototyping [46], and user testing [29] with end users and other key stakeholders. PSD principles from the PSD model [55] and BCTs from the Cross-Domain Taxonomy (version 1) by Michie et al [52] were combined and implemented into design features during several

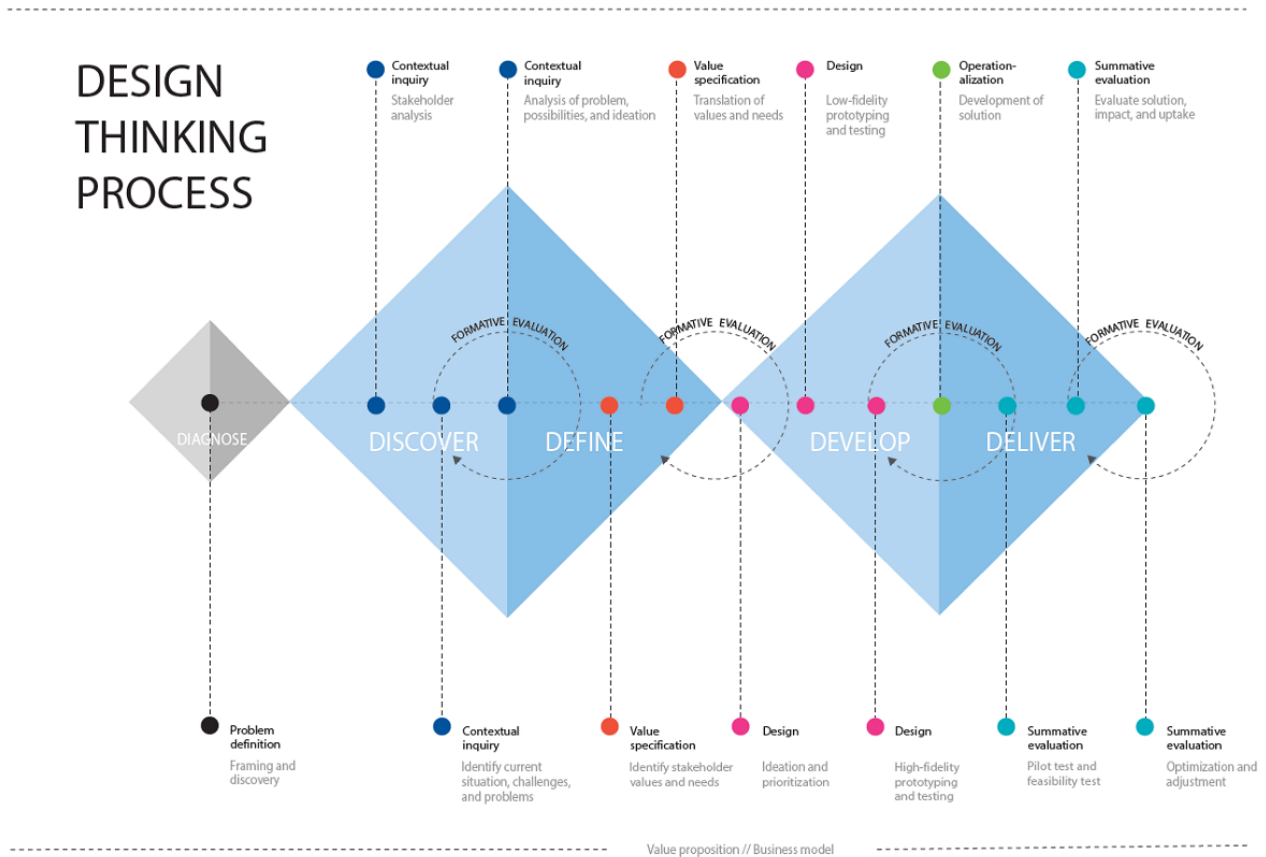
iterative cycles, providing ongoing information on how to address end user values and needs [32] (Figure 1).

The *Operationalization* phase [36] included Agile software development (ie, Scrum development sprints) [42] and testing of the digital intervention and focused on how the technology could be *delivered* and further improved to prepare for a feasibility pilot trial [29]. The PSD model [55] includes 28 individual PSD principles, which are categorized into 4 categories: *primary task support*, *dialogue support*, *system credibility support*, and *social support*. In comparison, the

Cross-Domain Taxonomy by Michie et al [52] comprises 93 distinct BCTs divided into 16 theory-independent clusters. The 2 frameworks applied to target behavior change [52,55] can facilitate standardized reporting of the PSD principles and BCTs embedded in a digital behavior change intervention.

In line with the iterative approach to eHealth development [36] and the Medical Research Council guidance for developing and evaluating complex interventions [54], the data collection, analysis, and results were intertwined throughout the design and development processes (Figure 2).

Figure 2. The Double Diamond [66] and the Center for eHealth Research and Disease Management Roadmap [36] combined: a design thinking process for eHealth design and development.



Multidisciplinary Research and Development Team

Digital intervention development involved a multidisciplinary research and development team and was led by the study principal investigator (PI; LSN), a clinical psychologist with health psychology specialization and long-standing experience with digital behavioral interventions. The multidisciplinary team (14/45, 31%) entailed key stakeholders identified through

stakeholder analysis [32] with diverse professional backgrounds and expertise, including researchers and clinicians in obesity and weight management, behavioral science, and eHealth; content editors; a digital designer; and software developers. In addition, a service designer was involved in facilitating service design workshops. Table 1 provides an overview of the multidisciplinary research and development team members, including their expertise.

Table 1. Overview of the multidisciplinary research and development team background and expertise (N=14).

Grouping	Total number, n	Obesity expertise, n	eHealth expertise, n	Behavioral and clinical health psychology expertise, n	Licensed health care providers ^a , n
Health care researchers	6	2	4	2	6
Content editors ^b	3	0	3	0	2
Design and software team	5	0	5	0	0

^aFor example, nurses, medical physicians, health psychologists, and physical therapists.

^b1 content editor acted as product owner during the development phase.

Recruitment of Study Participants: End Users and Other Key Stakeholders

End users were defined as people aged ≥18 years with a BMI of ≥30 kg/m² [67] before weight loss (ie, who had lost ≥8% of their body weight through a low-calorie diet or behavior change program) who were in need of support to prevent weight regain. People who met these criteria and were able to speak and read Norwegian were invited to participate in this study. Recruitment was conducted at 3 secondary or tertiary obesity research and treatment centers (ie, hospitals) in Norway through convenience sampling. In addition, end users who participated in a prior formative study [32] and a group of healthy volunteers were invited to participate in usability testing. To compensate for time spent and potential costs (eg, parking and transport), the study participants (ie, end users and healthy volunteers) received a gift certificate (ie, approximately US \$25 and US \$50 for individual testing and workshops, respectively).

Representatives of other key stakeholders identified during stakeholder analysis [32] (eg, health care providers and behavior

change and obesity experts) were recruited based on convenience sampling through the collaborating obesity research and treatment centers.

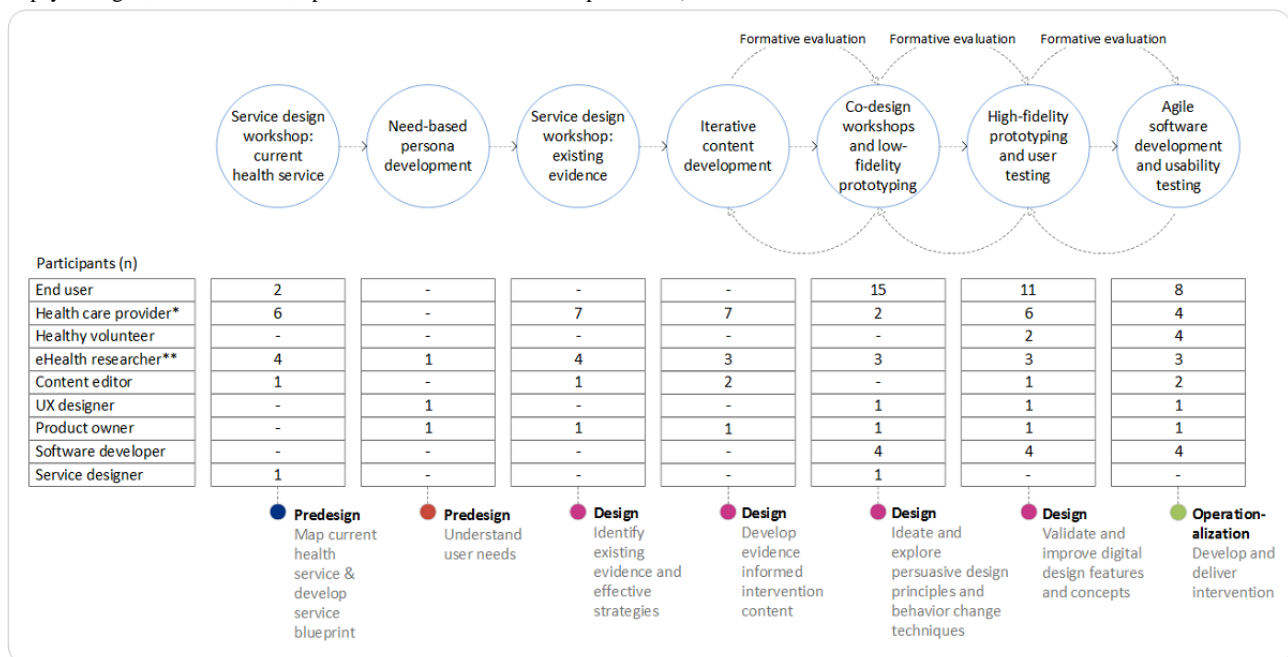
Ethics Approval and Informed Consent

This study was approved by the Hospital Privacy and Security Protection Committee (ie, institutional review board equivalent; approval number 2017/12702) at the Oslo University Hospital in Norway. All study participants (ie, end users and other key stakeholders) received written and oral study information and signed an informed consent form before participation.

Design and Development of the Digital Intervention

The development process started with 2 predesign activities to prepare for the *Design* and *Operationalization* phases to *develop* and *deliver* the digital behavior change intervention. An overview of the design and development cycles during the eCHANGE intervention, including activity participation, is shown in Figure 3.

Figure 3. Overview of the iterative development process of the eCHANGE intervention, including activity participation (ie, n is the number of participants), based on the Double Diamond Approach. See Asbjørnsen et al [32,35] for the previous research and development steps. Some participants (ie, 4 end users and 3 health care providers) participated in >1 activity.*Health care provider (eg, clinicians and researchers/experts in obesity management: medical doctor, clinical dietitian, exercise physiologist, physical therapist, geneticist, psychologist). **eHealth researcher (eg, registered nurse, clinical health psychologist, health scientist, specialist health education and promotion).



Pre-design Phase

Service Design Workshop Mapping Current Health Services

First, a service design workshop [41] was conducted with user representatives (2/45, 4%) and other key stakeholders (11/45, 24%; ie, health care providers, eHealth researchers, and content editors) to gain insights into the current health services offered and to optimize the fit between the context, needs, and technology to be developed. The first author (RAA; a researcher and eHealth expert) and a service designer served as workshop facilitators. To develop a service blueprint [68], visualizing the current health services from a user perspective (ie, including user needs and experiences), the workshop (3 hours) focused on mapping the current health services and user journey (ie, with sticky notes on a large whiteboard).

Need-Based Persona Development

To create an image of future end users and facilitate an understanding of their needs and challenges, need-based personas (ie, user profiles) reflecting a subset of the identified end user values and needs [32] were developed by a digital designer, product owner (MO; ie, Scrum) [42], and the first author (RAA), inspired by existing guidelines [69,70]. A total of 6 personas were created reflecting the target group aspects and containing information related to demographics, weight history, social and health-related factors, habits, and common everyday challenges and needs (Multimedia Appendix 2). The personas were used as a design tool to ideate, create, and reflect on prototypes during the design process, explore how the solution could be developed, and ensure that some of the identified key end user values and needs [32] were considered.

Design Phase

Service Design Workshop Related to Existing Evidence and Content Development

A second service design workshop [41] (4 hours) was performed with other key stakeholders (10/45, 22%; ie, health care providers and eHealth, obesity, and behavior change experts). The aim was to identify evidence-based strategies associated with successful weight loss maintenance, including but not limited to existing research from, for example, the US National Weight Control Registry, determinants of weight loss maintenance [12,71-74], national [75,76] and international guidelines for a healthy lifestyle and obesity management [77-80], and stakeholders' knowledge and experiences from weight management interventions. Relevant topics and themes to include in the content development to meet end user informational and educational support needs [32] were identified and facilitated (ie, using sticky notes and a whiteboard) by the first author (RAA) and a content editor.

Health care providers and researchers (10/45, 22%) from a wide range of disciplines with expert knowledge of obesity and weight management (eg, dietetics and nutrition, physical activity, physiology, medicine, and health psychology) contributed to

content development (ie, educational material and skills training). The experts were also consulted throughout the technology development process to secure high-quality, evidence-informed intervention content based on existing evidence and BCTs known to be effective and promising in supporting weight loss maintenance [32,35].

As indicated in Figure 3, the content development process underwent several iterations based on end user feedback during the design process. Content development was led by the PI (LSN), assisted by content editors, who were responsible for editing and optimizing all written material (ie, texts, including images, videos, and voice-to-texts) during the digital intervention development. The aim was to enhance the value of the tailored information, feedback messages, educational material, and skills training included in the intervention. This was done through close collaboration between multidisciplinary experts and the research team, coordinated and adapted by editors, to create positive and meaningful intervention content, grab end user attention, and motivate change. The language used (eg, readability, clarity, and tone of voice) was given particular attention during the content development process.

Co-design Workshops and Low-Fidelity Prototyping

A total of 4 co-design workshops [41] (ie, research and design cycles, 3 hours each) with end users (15/45, 33%) were organized to ideate and explore how to meet end user values and needs to successfully maintain weight loss in the long term [32]. The prototyping comprised small iterative steps to obtain information on how to combine and implement PSD principles and BCTs. Low-fidelity prototypes (ie, paper prototypes and simple Marvel sketches) [29,46] were developed by the digital designer, together with participants from the research and development team, based on the previously identified PSD principles, BCTs, high-level requirements, and suggested design features [32,35].

The co-design workshops, facilitated by a service designer and the first author (RAA), started with an ideation session [41] to stimulate creativity and develop ideas on how to support 3 of the identified key values (ie, happiness, social support, and motivation) [32]: "What makes you happy?" and "How can a Virtual Coach help you keep focus and stay motivated?"

Low-fidelity prototypes were used to explore and validate the following design features: (1) *personalized self-monitoring*, (2) *goal setting and planning*, (3) *smart feedback* (eg, praise, rewards, reminders, and suggestions), and (4) *shaping knowledge* (eg, education and skills training) through various human-centered design methods (Table 2 and Multimedia Appendix 3). The participants were encouraged to share, sketch, and reflect on ideas and paper prototypes during the co-design workshops. Sticky notes and participant drawings were used to collect user feedback on how to create an engaging and motivating self-management technology, supporting healthy behaviors and weight loss maintenance.

Table 2. Formative evaluation: detailed overview of methods in the *Design and Operationalization* phase.

Formative evaluation methods and procedures	Design and operationalization		
	Co-design and low-fidelity prototyping ^a	High-fidelity prototyping ^b	Agile software development ^b
A/B testing [43]			
Two versions of the design features (eg, horizontal vs vertical weight graph in relation to habits), were created, tested, and evaluated during to evaluate users' preferences and validate features/concepts	✓ ^c	✓	
Expert reviews [29]			
Operationalization and combination of PSD ^d principles [55] and BCTs ^e [52]	✓	✓	
Compliance with requirements for universal design, data protection by design and by default, and security guidelines (eg, web Content Accessibility Guidelines 2.0) [81-83]	✓	✓	✓
Scenario based tasks [29]			
Four specific scenarios and tasks: animated onboarding and goal setting, creating a Week Plan, personalization of the intervention, and selecting favorite knowledge and skills training; after evaluating, if tasks could be successfully completed, the facilitator asked questions about the user experience		✓	
Think-aloud technique [29]			
The participant could test the solution as they wished while sharing (ie, think aloud) what they did and why, accompanied by open-ended questions by the facilitator	✓	✓	✓
The Sauro System Usability Scale [84]			
A brief questionnaire about system usability with a 1 (strongly disagree) to 5 (strongly agree) Likert scale was performed when the participant was alone in the room		✓	✓

^aWorkshops facilitated by a service designer and/or first author.

^bIndividual sessions facilitated by the Scrum product owner.

^cIndicates which formative evaluation methods were applied.

^dPSD: persuasive system design.

^eBCT: behavior change technique.

High-Fidelity Prototyping and Usability Testing

During the next steps (Figure 3), the aim was to validate and improve the design features and intervention content at a high-fidelity level [29] during individual prototyping and usability testing sessions (45 minutes). A digital prototype of the intervention, developed using a web-based tool (ie, Marvel design platform), was presented on a smartphone to give a real-world look and feel and provide a certain degree of interaction (ie, gradually more realistic).

As presented in Table 2, a variety of formative evaluation methods were applied in combination to evaluate the high-fidelity prototypes and perform usability testing during the *Design* and *Operationalization* phases. The individual sessions with end users (12/45, 27%), health care providers (9/45, 20%), and healthy volunteers (4/45, 9%), facilitated by the product owner (MO), were voice/video recorded, whereas an observer (ie, digital designer, first author [RAA], or eHealth expert/content editor) collected notes. Some participants participated in >1 design and development cycle. The observer created a report with feedback and suggestions related to the design, content, and functionality, as well as additional input (eg, recurrent problems for, barriers to, and facilitators of use).

Stakeholders from the multidisciplinary research and development team (ie, eHealth researcher, digital designer, and

software developers) contributed to prototype development based on participant feedback, focusing on graphical and conceptual design and performing expert reviews during the design and development process. Table 2 provides an overview of the formative evaluation methods applied, and Multimedia Appendix 3 includes additional details about participants involved.

Operationalization Phase

Agile Software Development

The *Operationalization* phase centered around the operationalization of design concepts through short development cycles based on Agile and test-driven development principles [43] to deliver a minimal viable product (MVP) for pilot testing in the real world (ie, ClinicalTrials.gov NCT04537988). The technology was developed during four incremental development sprints (ie, Agile software development and Scrum) [85] and is divided into (1) *habit tracking and smart feedback* (including rewards; ie, *Week Plan*), (2) *registrations and self-monitoring* (ie, *My Overview*), (3) *virtual coach and tips* (ie, *Virtual Coach and Smart, Tailored Feedback*), and (4) *animated effects, knowledge, and strategies* (ie, *Knowledge and Skills Training*). To optimize the efficiency of the development process, some of the iterative design and development cycles (ie, sprints) ran simultaneously.

To secure flexibility and quick adaptation of features based on the feedback collected during continuous validation and usability testing, the development team deployed a web-based software collaboration tool (ie, on the GitHub development platform) and daily stand-up meetings (ie, based on the Scrum methodology) [85]. The high-fidelity prototypes (ie, in Marvel), together with accessibility guidelines, personas, specified user stories (ie, functional requirements), and technical stories (ie, nonfunctional requirements), served as bases for the software development process. The user stories reflected the user needs and requirements for a desired feature and were intended to help the development team understand end user needs in relation to the system and its context. The user stories were written in an informal way from the user's perspective: "As a [description of user], I want [functionality], so that [benefit]." Examples of user stories included the following: "as a user, I want to create my own plan, so that I can choose which habits I want to work on" or "as a user, I want to receive reminders, so that I do not forget to work on my habits."

To maximize value and set the direction for software development (ie, Agile Scrum team), the product owner (MO) ensured that the product backlog was up to date (ie, prioritized list of requirements and acceptance criteria) and secured, along with the design and development team, rapid delivery of implemented features, and high-quality software based on user needs and requirements.

Data Analysis

During the iterative intervention development (Figure 3), data collection and analysis provided ongoing information on how to improve the technology (eg, intervention content and operationalization of the design features). The prototypes evolved during co-design, rapid prototyping, and usability testing, where findings from one design cycle served as input to the next and set the direction for creation and validation of design features.

Rapid analysis [86] was applied during the design and development process to ensure that the collected data provided quick and thorough input to optimize the digital intervention (eg, actionable suggestions and specification of requirements through user stories). This included the structuring and summarizing of notes, including illustrative quotes and voice/video recordings from usability testing, into themes related to the design features to elicit input on how to combine and implement the PSD principles [55] and BCTs [52] into design features to support end user needs and preferences. In addition, a more in-depth directed content analysis [87] was performed to secure evidence-informed development and ensure that no themes were missed. This included the coding of study participants' feedback and suggestions into predefined categories with respect to design, content, and functionality. The feedback and findings, including inconsistencies and conflicting needs, were first reviewed and discussed by the PI (LSN), 2 researchers (RAA and MLS), the product owner (MO), the digital designer, and/or a content editor to validate that the intervention content and design features (ie, PSD principles and BCTs) matched the identified end user values and needs. Thereafter, the findings were discussed (until consensus was reached) with participants

from the multidisciplinary research and development team for continuous evaluation, prioritization, and improvement of features before the next iteration was conducted.

The product backlog reflected the prioritizations and decisions made by the research and development team. This was done to ensure that the selected design features and development of the intervention (ie, MVP) were in line with end user needs and preferences and that findings from previous research [32,35], evidence-informed knowledge, and feasibility considerations were taken into account before technical adaptations and updates were executed.

Technical Architecture

The smartphone-based eCHANGE intervention app was designed and distributed as a native app through official app stores for iOS and Android. Web technology in a Cordova container was used, and when in use, all information is stored locally and encrypted with the Advanced Encryption Standard algorithm in Galois/counter mode before being written to the local device file system.

The first time the app runs, a 256-bit encryption key is generated. Between app invocations, an Advanced Encryption Standard key-wrapped algorithm with a (wrapping) key acquired from the user's personal identification number is used and stored on the device keychain (ie, based on an existing technology platform) [88]. The keychain ensures that the mobile device is protected with a 4-digit personal identification number code or optional biometric authentication (ie, face recognition or fingerprint).

When in use, log data (ie, system use, including navigation, frequency of use, and use of functionalities) and self-monitoring data are sent through an encrypted channel to a secure server (ie, Services for Sensitive Data, University of Oslo) for future analysis and summative evaluation. If interested, the user may elect to import relevant personal data (eg, number of steps taken per day and weight measurements) from Apple Health or Google Fit on their device; however, this integration is neither required by the research team nor necessary to realize the full potential of the app.

Security and Privacy Considerations

The eCHANGE intervention was developed by the Department of Digital Health Research at Oslo University Hospital in Norway in line with the national and international privacy and security standards and regulations (eg, Norwegian Digitalization Agency and European General Data Protection Regulations of 2018) [81-83]. A legal and privacy declaration was included as part of the information about the app in the settings functions in accordance with the existing requirements. The procedures applied in relation to the data protection impact assessment and risk assessment analysis of the technical solution were approved by the institution's Department of Information Safety (approved in June 2020).

Results

Overview

This section focuses on the design operationalization of the incorporated PSD principles and BCTs into design features and main intervention components to address the following research question: *how can an evidence-informed digital behavioral change intervention, combining and implementing PSD principles and BCTs into design features, be designed and developed to support end user values and needs for long-term weight loss maintenance?*

To meet the diverse and dynamically changing needs of end users (ie, people aiming to maintain weight after weight loss) [32], the eCHANGE intervention was developed as a personalized self-management intervention and delivered as a smartphone-based app for flexible and easily available weight loss maintenance support.

Participants

A total of 45 end users and other key stakeholders participated in intervention development, including 13 (29%) external stakeholders (ie, n=9, health care providers, including 1 health care manager and 4 healthy volunteers), 17 (38%) prospective end users, 14 (31%) stakeholders from the multidisciplinary

research and development team, and 1 (2%) service designer. As presented in [Multimedia Appendix 4](#), most end users were female (12/17, 71%), and the median age was 48 (range 30-63) years.

Combining and Implementing PSD Principles and BCTs Into Design Features to Support End User Values and Needs for Long-term Weight Loss Maintenance

Identified Design Features

The various research and development activities performed in the study explored how a digital intervention can support identified key end user values (ie, *personalized care, feel supported, positive self-image, health, happiness, motivation, autonomy, and self-management*) [32] to maintain weight after weight loss. The study co-design and prototype validation resulted in the selection of 17 design features for (sustainable) behavior change to be included in the digital intervention (MVP). Further details are provided in [Table 3](#).

The design features and main components of the digital intervention described in the following sections provide insight into how PSD principles and BCTs were combined and implemented to support end user values and needs for long-term weight loss maintenance.

Table 3. Identified design features and main components to support key end user values for long-term weight loss maintenance.

Main components and design features	Key end user values [32]							
	V1 ^a	V2 ^b	V3 ^c	V4 ^d	V5 ^e	V6 ^f	V7 ^g	V8 ^h
(A) Animated onboarding	✓ ⁱ						✓	✓
Week Plan^j								
(B) Behavioral planning and goal setting (eg, action and coping planning)	✓	✓				✓	✓	✓
(C) Motivational exercise and realistic goal setting		✓	✓			✓	✓	✓
(D) Habit rehearsal and tracking	✓	✓	✓				✓	✓
My Overview								
(E) Personalized self-monitoring	✓	✓	✓	✓	✓	✓	✓	✓
(F) Goal setting target outcome	✓					✓	✓	✓
(G) Automatic integration of data	✓			✓	✓	✓		✓
(H) Visualization of target behavior in relation to target outcome	✓	✓	✓	✓	✓	✓	✓	✓
Knowledge and Skills								
(I) Educational material and information	✓	✓	✓	✓	✓	✓	✓	✓
(J) Cognitive and motivational exercises	✓	✓	✓	✓	✓	✓	✓	✓
(K) My favorites	✓					✓	✓	✓
Virtual Coach and Smart Feedback System								
(L) Virtual coach	✓	✓	✓	✓	✓	✓	✓	✓
(M) Animated nudging elements		✓			✓	✓		
(N) Praise	✓	✓	✓	✓	✓	✓	✓	✓
(O) Rewards	✓		✓		✓	✓		
(P) Reminders	✓					✓		✓
(Q) Suggestions		✓		✓		✓	✓	✓

^aV1: personalized care.

^bV2: feel supported.

^cV3: positive self-image.

^dV4: health.

^eV5: happiness.

^fV6: motivation.

^gV7: autonomy.

^hV8: self-management.

ⁱIndicates the design features identified to support the values of end users aiming to maintain weight after weight loss.

^jIndicates the main intervention components.

Design Features and Main Intervention Components

The eCHANGE Intervention

The development of the eCHANGE intervention, incorporating the 17 design features presented in Table 3, resulted in an adaptive, interactive, and interconnected concept with four main components, as shown in Figure 4: (1) *Week Plan*, (2) *My Overview*, (3) *Knowledge and Skills*, and (4) *Virtual Coach and Smart, Tailored Feedback*.

On the basis of end user input wishing for a user-friendly, motivating, and personal intervention *that fits me*, an animated onboarding introduction (ie, design feature A, Table 3) was

created to present the main features of the eCHANGE app. The onboarding feature could then be used to record baseline data (eg, current weight and weight maintenance goal) and tailor the intervention (eg, motivating messages and suggestions) to individual goals and needs. A general settings function for individual system preferences was also created to meet the end users' individual preferences and needs for self-management support. This function evolved based on user and other key stakeholder feedback (ie, including usability experts from the multidisciplinary team) and included a range of personalization options (eg, frequency of reminders, type of feedback messages, automatic exchange of data, and dark or light mode).

Figure 4. eCHANGE conceptual design and main intervention components to support end user values and needs.



Week Plan

During the co-design workshops and prototyping sessions, end users and health care providers emphasized the need for technology supporting planning and adherence to healthy lifestyle habits in their pursuit of maintaining weight and focusing on health and well-being (ie, not only weight). This included strategies on how to manage *high risk situations* (eg, situations with availability of tempting foods/snacks) and help with *impulse control* (eg, resistance to impulsive behavior such as comfort eating) to overcome lapses and prevent relapse into previous behaviors. One of the participants stated the following:

I would like to create a plan and choose which habits and goals to work on, kind of like a calendar. [End user]

A personal *Week Plan* was subsequently created to meet individual preferences and needs for healthy lifestyle changes to maintain weight. [Figure 5](#) shows the eCHANGE *Week Plan* screenshot examples and included design features. The *Week Plan* contained options to select habits and strategies to work on from the following 4 categories: eating habits (eg, eating breakfast, planning meals, and healthy meals/snacking), physical activity habits (eg, daily walking goal), well-being (eg, sleep,

stress management, and mindfulness), and self-regulation strategies (eg, problem solving, if-then plans, and back-on-track plan) (ie, design feature B, [Figure 5](#)), all associated with long-term weight loss maintenance [8,12,23,72,73].

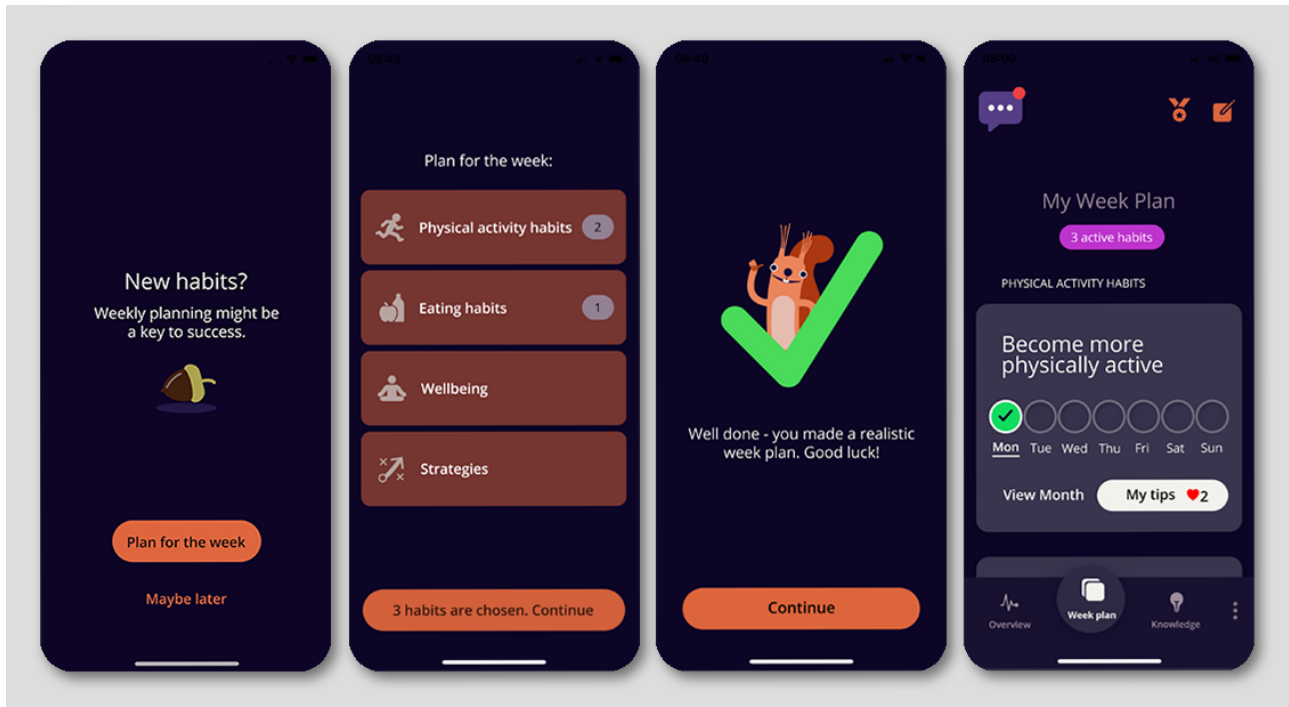
Graded tasks (eg, subgoals or easy-to-reach targets) could be selected by the end user to facilitate self-efficacy and self-belief, as well as the adoption of physical activity and healthy eating habits. To support flexibility in planning, the *Week Plan* included an option for user-initiated changes whenever needed.

Information about health effects and *My favorite tips* (eg, introduction of environmental cues and restructuring of the physical environment) were also added to the habits and strategies presented in the *Week Plan*. This was done based on user feedback, aiming to provide knowledge and practical tips on how to adopt and maintain healthy habits as part of their daily routine. On the basis of suggestions from health care providers, a motivational exercise inspired by motivational interviewing techniques [89] was also incorporated to encourage realistic goal setting and stimulate motivation and self-efficacy during planning (design feature C, [Figure 5](#)). Building on previous formative results [32], as well as input from end users and eHealth experts, engaging and motivating design elements

were added during the *Week Plan* design process. These design elements included tracking and monitoring of self-selected habits (ie, by checking the box in the personal plan; design feature D, [Figure 5](#)), animated prompts when tracking, weekly

and monthly rewards related to personal targets, and positive feedback messages to stimulate motivation and adherence to healthy behaviors.

Figure 5. Screenshot of the eCHANGE program (ie, dark mode). Week Plan included the following design features: (B) behavioral planning, (C) motivational exercise and realistic goal setting, and (D) habit rehearsal and tracking.



My Overview

Participating end users highlighted the need for an easy overview of the progress and performance of behavioral goals related to their weight maintenance goal, with the possibility of automatic integration of data (eg, from existing health apps and wearables). One of the participants stated the following:

I would like to have overview of my data in one place, to understand what works and does not work in order to maintain weight [End user]

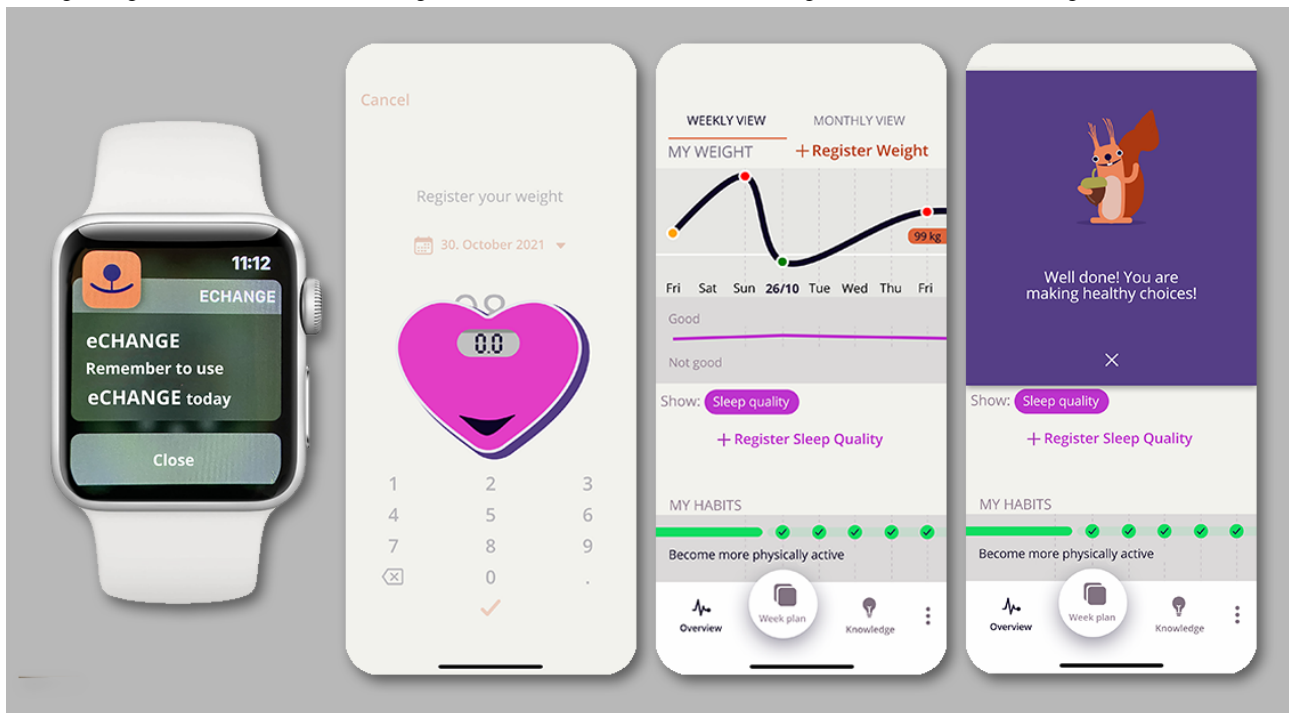
Despite wishing for ways of monitoring progress, some end users expressed that they did not want to be *forced* by the system to register their weight. In contrast, health care providers emphasized the importance of daily or, at minimum, weekly self-monitoring of weight to encourage self-regulation and prevent regain. On the basis of this feedback, the component *My Overview* was created to support self-regulation and facilitate personalized self-monitoring. [Figure 6](#) shows the eCHANGE *My Overview* screenshot examples and included design features.

In response to participant feedback, the possibility of automatically transferring data (ie, weight and steps) from existing health apps (ie, Apple Health and Google Fit) was included in *My Overview* to simplify self-monitoring and

facilitate awareness and engagement. Health care providers (ie, obesity experts) also suggested a *traffic light* system [90] based on 3 color zones, as illustrated in [Figure 6](#), design feature H, to provide visual and tailored feedback based on which *weight zone* users are in (ie, in relation to their target weight). The "green zone" was defined as <1.5 kg (ie, <3.3 lbs) above the target weight (eg, indicating to be "on track"), the "yellow zone" was when the weight increase was 1.5-3 kg (ie, 3.3-6.6 lbs) above target weight, and the "red zone" was an increase of >3 kg (ie, >6.6 lbs) above target weight.

In line with feedback from users not wanting to be forced to register weight, some end users also reported not wanting to focus on, or seeing, their body weight *all the time* and stated that they wished to be able to use the app in public spaces or show their progress to family and friends, without revealing their actual weight. Therefore, *My Overview* also included options for the user to choose when to register weight, whether to visualize actual weight or discrepancy from target weight, and a *hide the weight* option by clicking on their current weight. As end users highlighted holistic self-monitoring as an important feature, the possibility of monitoring self-selected habits, physical activity (ie, steps), stress, sleep, and mood over time was also incorporated in *My Overview*.

Figure 6. Screenshot eCHANGE program (ie, light mode). My Overview included the following design features: (E) personalized self-monitoring, (F) goal setting of target outcome, (G) automatic integration of data, and (H) visualization of target behavior in relation to target outcome.



Knowledge and Skills

Several end users and health care providers in the study expressed the need for technology with trustworthy information about weight loss maintenance, including trustworthy (ie, evidence-based) information about strategies and skills to support and improve end users' competence, autonomous motivation, self-regulation, and the ability to prevent weight regain. In response to this input, as well as previously identified end user values and informational support needs [32], a *Knowledge and Skills* component was created.

Content for this component was identified through end user and obesity specialist feedback. For example, the service design workshop with obesity management and behavior change experts identified 15 weight loss maintenance–related topics that were important to include in a weight loss maintenance–specific *Knowledge and Skills* section.

During the co-design sessions, end users also emphasized the need for educational material and information to be provided in an appropriate and understandable language, with brief textual information supported by images or videos and an audio option for listening rather than reading. This was implemented through iterative stakeholder testing. On the basis of input from end

users, a *My favorite* option for the included information or exercises could also be chosen to facilitate easy access to relevant content based on individual preferences. Figure 7 shows eCHANGE *Knowledge and Skills* screenshot examples and the included design features.

In the co-design and prototyping sessions with end users and other key stakeholders, health and well-being, happiness, feeling of control and mastery, and motivation for long-term change were also identified. One of the participating obesity experts stated the following:

Many people lose faith in their capability to maintain weight [Obesity expert]

To support autonomous motivation, self-belief, positive body image, self-efficacy, and self-regulation skills to prevent regain, experts on obesity and behavior change also identified 25 cognitive and motivational exercises to be included in the intervention. Theory-based exercises in the final eCHANGE intervention were anchored in BCTs identified as important to address sustainable behavior change to successfully maintain weight [32,35] and recognized motivational, self-regulation, and cognitive behavioral theories [89,91-96]. Table 4 provides an overview of the topics available in the eCHANGE *Knowledge and Skills* section.

Figure 7. Screenshot eCHANGE program. Knowledge and Skills included the following design features: (I) educational material and information, (J) cognitive and motivational exercises, and (K) my favorites.

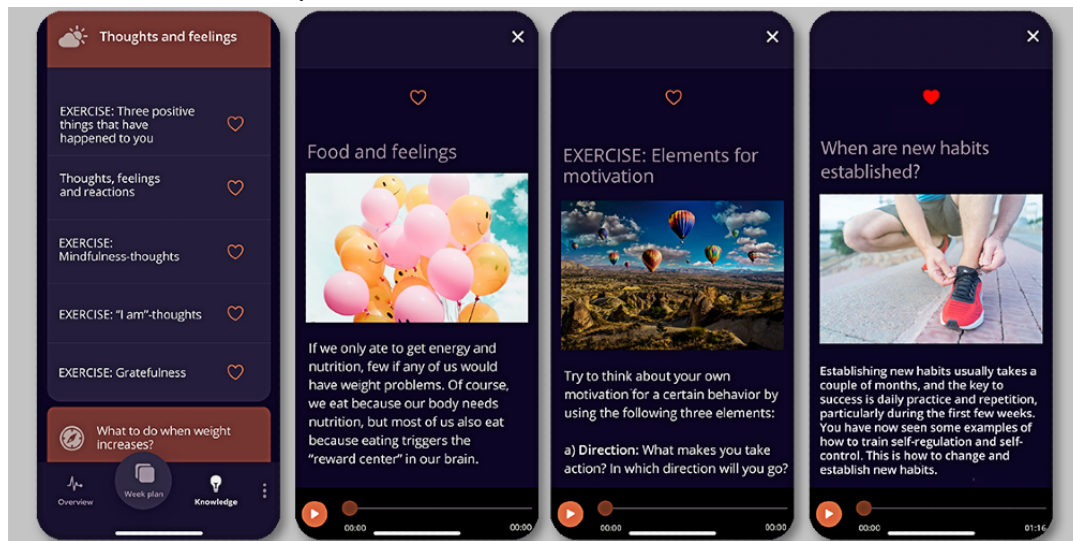


Table 4. Overview over topics and content included in the eCHANGE *Knowledge and Skills* section.

Topic number	Topic	Content
1	Introduction	Introduction to the intervention program, main components, and general information about weight loss maintenance
2	Adaptive thermogenesis and energy balance	Information about body/physiological processes and challenges to maintain weight, including strategies to prevent weight regain
3	What is important to me?	Information and exploration of values; self-image, personal role models, identity, and thought patterns; value prioritization and life goals
4	How to change habits	About being in charge of own life, the nature of habits, awareness, behavioral patterns, and habit substitution; thoughts and behavior change and implementation of new habits
5	Becoming friends with the scale	Addresses the importance of self-monitoring and self-awareness for behavior change; thought patterns, positive self-talk, and self-confidence
6	Goal setting, planning, and problem solving	Defining realistic goals, regulation, and planning of healthy habits; relapse prevention and if-then plans and self-monitoring toward a personal goal
7	Motivation	Identity and values, internal drivers of behavior, types and factors of motivation, self-belief and behavior, thought patterns and self-belief, motivation, and relatedness
8	Food and drinks	Healthy diet and health effects; food and emotions/stress; healthy behaviors and health behavior change, awareness, habits and routines, nutrition, and healthy eating strategies
9	Physical activity	Physical activity and weight loss maintenance, barriers or physical challenges, strategies on how to incorporate physical activity into daily life, and training/exercise suggestions
10	Sleep	Information about circadian rhythm and sleep, importance of health, and quality of sleep; improvement of sleeping routines and health effects
11	Communication	Communication and your surroundings, body language, self-image, and positive self-talk
12	Social support	Types of social support and skills to strengthen social support systems; peer support; social and environmental cues for healthy and unhealthy habits and stimulus control
13	Thoughts, feelings, and stress	Relationship between thoughts and feelings, thoughts and stress, regulation of emotions and thoughts, thought reframing, self-efficacy, and positive self-image/self-esteem and body image
14	Weight maintenance and weight regain	Strategies for successful weight loss maintenance; traffic light system and weight zones; weight regain and causes; skills for self-regulation, problem solving, and relapse prevention
15	Mindfulness and relaxation	Introduction to mindfulness; practice self-compassion and strategies for stress management, including relaxation/mindfulness exercises to improve health and well-being

Virtual Coach and Smart Tailored Feedback

During the ideation and co-design sessions with end users, the need for positive, tailored, *just-in-time support* to reinforce healthy behaviors and motivation *to keep up and stay focused* was highlighted. This resulted in the co-design of a *virtual coach*, designed as an animated squirrel, as well as a *smart feedback system* developed to provide decision support and stimulate motivation and adherence to healthy behaviors. Although some end users expressed a wish to create their own personalized *virtual coach*, the essential factors, according to user feedback, involved content and expressions. One of the users stated the following:

It is not important how it (ie, the buddy) looks, but what it says and what it does. [End user]

The *virtual coach* was co-created with end users and other key stakeholders to facilitate gradual engagement and guide the user through the first *setup* of a personal *Week Plan* and *My Overview*. Participating obesity and behavior change experts stated that the *virtual coach* should also provide feedback in line with professional coach advice. This led to the incorporation of feedback messages developed by health care professionals, with coaching techniques in accordance with motivational interviewing [89]. Figure 8 shows the eCHANGE *Virtual Coach* and *Smart Feedback System* screenshot examples and included design features.

To provide *just-in-time support*, the automated, smart feedback system was developed based on real-time self-monitoring data, dynamically adapting to the user based on a set of rules (ie, preference based and data-driven algorithms). The feedback system, developed by software experts in collaboration with

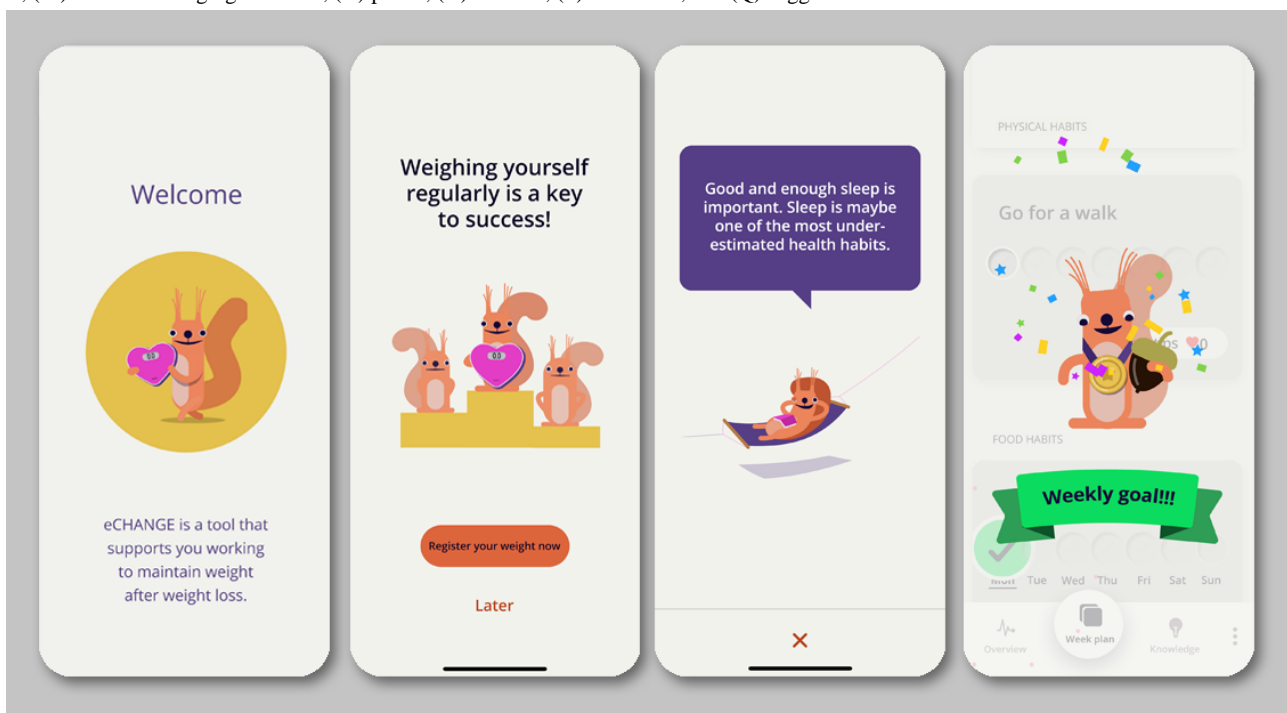
obesity and behavior change experts, allowed tailoring of the intervention to individual progress. This included positive feedback on the performance of the behavior, close to target behavior; healthy lifestyle suggestions; information about health benefits of healthy lifestyle behaviors; suggestions to make healthy choices based on current weight zone (ie, green zone, yellow zone, and red zone); and help to get *back on track* (eg, when in the yellow or red weight zone). Multimedia Appendix 5 provides examples of feedback messages.

End users also reported preferring a *virtual coach* that would support habit formation through health-focused suggestions on how to maintain weight, prompt daily rehearsal of target behavior, and bring a sense of joy or happiness by suggesting new habits. Several animated nudging elements were integrated into the behavioral design during the high-fidelity prototyping and usability testing (Figure 8, design feature M) to support these preferences, promote healthy behaviors, and make the intervention desirable. When testing the intervention following this incorporation, one of the end users reported the following:

Opening the app puts me in a good mood [End user]

The participating end users and health care personnel also suggested that the *virtual coach* should provide reminders that could stimulate adherence to the *Week Plan* and self-monitoring of weight, as well as weekly and monthly rewards to highlight goal achievement, as presented in Figure 8, design features O and P. In support of these suggestions, previous findings [32,35], and input from participating behavior change experts, encouragement to apply positive self-talk or choose self-selected rewards in line with individual values were incorporated to facilitate autonomous motivation and ongoing behavior change, without threatening autonomy.

Figure 8. Screenshot eCHANGE program. Virtual Coach and Smart Tailored Feedback System included the following design features: (L) virtual coach, (M) animated nudging elements, (N) praise, (O) rewards, (P) reminders, and (Q) suggestions.



PSD Principles and BCTs Implemented in the eCHANGE Intervention

To support weight loss maintenance values and the needs of end users, PSD principles from the 4 categories of the PSD model by Oinas-Kukkonen [55], such as *tailoring*, *personalization*, *self-monitoring*, *reminders*, *rewards*, *rehearsal*, *praise*, and *suggestions*, were implemented in the eCHANGE intervention. The PSD principles were combined with BCTs from 15 of the 16 BCT clusters from the Michie et al [52] taxonomy of behavior change techniques, such as *goals and planning*, *feedback and monitoring*, *social support*, *repetition*

and substitution, *shaping knowledge*, *natural consequences*, *associations*, *antecedents*, *identity*, and *self-belief*.

The PSD principles of *personalization* and *tailoring* were incorporated into all main components to tailor the intervention to individual preferences and end user needs. For a description of the design features in the eCHANGE intervention and an overview of the design features, including PSD principles and BCTs combined and implemented, refer to [Textbox 1](#) and [Multimedia Appendix 6](#) [32,35,52,55]. Additional information about the formative design results from the system usability check (eg, System Usability Scale score) can be found in [Multimedia Appendix 7](#) [84].

Textbox 1. Design features description to support sustainable behavior change and weight loss maintenance.

(A) Animated onboarding

Animated onboarding introduces the app and guidance to create a personal *Week Plan* and *My Overview* to shape the future interaction (eg, how the content is delivered and visualized)

(B) Behavioral planning and goal setting

Supports creation of action and coping plans with self-selected healthy habits and goals, preplanning for potential barriers, and prevention of relapse (eg, weekend temptations and back-on-track strategies).

(C) Motivational exercise and realistic goal setting

Motivational exercise supports realistic goal setting and creation of a feasible *Week Plan*; the exercise (eg, self-evaluation bar 1-10) facilitates reflection and commitment to the plan

(D) Habit rehearsal and tracking

The *Week Plan* includes a tracking tool to follow up on self-selected healthy habits in a weekly or monthly overview (ie, calendar function), monitor progress, and support rehearsal and adherence to the plan

(E) Personalized self-monitoring

Personalized self-monitoring allows for registration of body weight, physical activity (ie, steps), perceived mood, stress, and sleep over time; visualized in a weekly or monthly overview; historical data is available

(F) Goal setting target outcome

Goal setting allows the user to set an outcome goal (ie, weight target/weight maintenance goal); actual weight or discrepancy from target weight (eg, +3 kg or -3 kg) is visualized in *My Overview*

(G) Automatic integration of data

Automatic integration of data is available for activity tracking (ie, steps) and body weight (ie, through Apple Health and Google Fit)

(H) Visualization of target behavior in relation to target outcome

A *personal visualization* of the target behavior(s) (ie, chosen habits) in relation to target outcome (ie, weight) over time in a weekly or monthly overview through graphs and icons in relation to weight zones (ie, green, yellow, and red); provides means for understanding the link between cause and effect of behavior and outcome to support awareness, self-reflection, and self-regulation; a progress bar (ie, goal gradient) related to each habit, visualization of goal progress and adherence to the plan and behavioral performance in relation to individual targets

(I) Educational material and information

Educational material and information through 15 topics related to sustainable weight loss maintenance and behavior change; provided through text or audio and videos

(J) Cognitive and motivational exercises

Skill training through 25 cognitive behavioral and motivational exercises to support skills related to behaviors, thoughts, and emotions and improve self-belief (eg, focus on past success and positive self-talk) and continued motivation for sustainable behavior change (eg, identity and personal values)

(K) My favorites

A general personalization feature where the user can mark and view only *My favorite* tips, skills training, knowledge, exercise, and/or strategies for easy access to personalized content and decision support

(L) Virtual coach

A *virtual coach* (ie, animated coach/buddy) provides automated, tailored (decision) support (ie, smart feedback—a data- and preference-driven algorithm enables smart feedback and tailoring of the intervention), including motivating messages, prompting of weight maintenance strategies, information about health effects, and self-reward/self-praise when reaching goals or performing target behavior. The virtual coach adopts a social, supportive role (eg, motivating interviewing techniques) and provides real-time progress/performance feedback related to health maintenance behaviors based on outcome data or weight zones (ie, green, yellow, or red zone in relation to target weight), physical activity data, habit tracking/behavior self-monitoring, and user data from the past 30 days; in the general settings function, type of feedback messages can be selected based on personal preferences

(M) Animated nudging elements (eg, prompts/cues)

Animated elements were provided to prompt, encourage, and positively reinforce healthy behaviors and decisions to reach target goals/desired behavior through enjoyable and surprising animated elements, such as a “heart scale” pop-up to encourage weight registration, animated effects (ie, firework/sparks) to stimulate adherence, and animated prompts close to target behavior or to elicit/trigger healthy behaviors

(N) Praise: positive feedback

Praise is provided through positive, tailored feedback messages close to target behavior (ie, real time) when reaching individual goals to recognize efforts and success and, unexpectedly, to stimulate motivation to sustain a healthy lifestyle

(O) Rewards

Rewards (eg, medal and confetti) are provided when reaching self - selected healthy habit targets to highlight goal achievement, facilitate engagement, and positively reinforce progress (eg, weekly/monthly reward)

(P) Reminders

Reminders through “pop - up messages” on a mobile device/compatible smartwatch to facilitate engagement and adherence (eg, behavioral practice and weight registration); personal frequency and type of reminder choice

(Q) Suggestions

Suggestions provided by the virtual coach or through animated prompts/cues to support healthy lifestyle habits (eg, suggestions of healthy habits and practical strategies in everyday life to keep weight off)

Discussion

Principal Findings

This study provides insights into the design and development of a digital behavior change intervention called eCHANGE, which aims to combine and implement PSD principles and BCTs into design features to support end user values and needs for long-term weight loss maintenance. The results revealed specific design features for sustainable health behavior change to prevent weight regain, combining PSD principles and BCTs, as well as how these design features could be operationalized into core components during the design and development of the digital intervention.

Combining and Implementing PSD Principles and BCTs Into Design Features to Support End User Values and Needs for Long-term Weight Loss Maintenance

On the basis of participant feedback, 17 design features were identified in this study (Table 3) to support 8 previously identified key end user values for weight loss maintenance [32]. During the co-design and prototype sessions, the values of *self-management*, *personalized care*, and *motivation* received the most attention and feedback from end users. The identified design features were implemented in the digital intervention through 4 interconnected main components: *Week Plan*; *My Overview*; *Knowledge and Skills*; and a *Virtual Coach and Smart, Tailored Feedback System*.

The findings indicate that to support the identified, interconnected end user values and needs [32], digital weight loss maintenance interventions should include design features that focus on health and well-being (ie, not only weight); facilitate the generation of habitual behavior, self-regulation, autonomous motivation, knowledge, and skills; and provide positive, tailored support to maintain weight after weight loss in the long term. Existing research has pointed toward a need for such an approach [32,35] and that a combination of PSD, BCTs, and behavior change theories might facilitate the design of effective technology-based tools and strategies for behavioral obesity interventions [35,58]. However, to the best of our knowledge, this is the first study to show how PSD principles and BCTs can be translated into design features to support end user values and needs to maintain weight in the long term. In line with existing research [33,97], the findings indicate a need for a shift in goal focus from *weight* to *health and well-being*. This may suggest that technologies for digital weight loss maintenance should aim to support healthy lifestyles and

sustained motivation in line with the self-determination theory, supporting psychological needs such as competence, relatedness, and autonomy [26,92,95]. Digital interventions incorporating features supporting habit formation in line with self-determined goals, individual values, and identity, as well as focusing on meaningful areas such as improved health, might also enhance a sense of purpose and facilitate long-term health behavior changes and weight outcomes [17,26,32,98-101]. The results also indicate that the application of PSD principles from the *primary support*, *dialogue support*, *social support*, and *credibility support* categories might be required to aid weight loss maintenance and that the application of PSD principles such as *personalization* and *tailoring* of design features is important to match individual challenges, goals, and key values.

The eCHANGE intervention was developed as a personalized (adaptive) digital intervention aimed at providing self-management support for long-term weight maintenance. The results illustrate how PSD principles can be combined with BCTs that have been identified [32,35] as effective and promising in supporting weight loss maintenance and how they can be integrated into a digital intervention based on multidisciplinary stakeholder feedback. Existing research has indicated that BCT combinations targeting motivation and persistence in health-promoting interventions might increase the chances of successful health behavior change [102]. A systematic review identifying active ingredients in complex behavioral interventions for adults with obesity also indicated that the inclusion of BCTs in interventions could be beneficial, facilitating assorted phases of the behavior change process [53]. As such, the adaptive intervention format of eCHANGE can potentially support the behavior change required to maintain weight, allowing the user to choose between a number of BCTs depending on individual needs.

Development of healthy habits and adjustment or *breaking* of less healthy habits are required to optimize and maintain the health benefits of weight loss, as well as maintain new weight, over time [103,104]. Therefore, self-regulation strategies and skills are essential to enhance when new physically active lifestyles or healthy eating patterns are not fully automated and are likely vital to maintain healthy behaviors and not relapse into previous habits [26,105]. The findings from this study show that targeting holistic aspects (eg, cognitive, emotional, social, and behavioral) of behavior change [103] during the design of digital technologies may be essential when aiming to deal with the often multifaceted challenges and needs of weight loss maintenance [23,32,98]. Informed by the self-determination and self-regulation theories [91,92,95], some of the features identified in this study focus on the generation of habitual

behaviors through recognized self-regulation and habit formation techniques. The application of habit-forming techniques could free cognitive capacity to facilitate engagement in desired behaviors that may help adopt health-related behaviors to maintain weight loss in the long term [104,106,107]. The incorporation of BCTs in digital interventions, such as self-regulatory strategies to bridge the intention-behavior gap (eg, *if-then plans/problem solving*), habit-related techniques (eg, *graded tasks*), and self-belief (eg, *focus on past success, self-talk*), could also be important for targeting specific values (eg, self-management, motivation, and positive self-image) or mechanisms of action (eg, behavioral regulation and beliefs about capability) that may facilitate continued behavior change and weight gain prevention [15,25,32,104,108-112]. A literature review aimed at identifying links between BCTs and mechanisms of action [112] indicated that some of the core BCTs incorporated in the eCHANGE intervention could affect behavior change mechanisms (eg, *beliefs about capabilities, behavioral regulation, and motivation*), which is in line with identified weight loss maintenance values and needs [32]. This underlines the importance of keeping in mind the links among intervention content, design features, and the values and needs of the target group when designing and developing digital interventions for weight loss maintenance.

Human behavior is affected and shaped by a range of individual factors, including cognitive, psychological, biological, physical, and emotional factors; habits; values; motivational and demographic factors; and external factors such as environmental, cultural, social, and physical contexts in which behaviors occur [113]. Therefore, when designing digital technologies for sustainable behavior change, a fit between technological, human, and contextual factors is required [36,114]. In addition, the success or failure of digital interventions subsequently depends on whether individual end user needs are met [32,115]. The development of technology-supported programs for weight loss maintenance based on theory, evidence, and person-based approaches is gradually receiving increased attention [116-119]. The eCHANGE intervention's value-based approach to eHealth development, in which PSD principles and BCTs identified to match end user values and needs [32] were gradually embedded into the digital technology through the design features, represents a major novelty. However, how persuasive and behavior change strategies are operationalized into design features may affect the usefulness and effectiveness of technology in terms of health outcome improvements. This points to potential issues, as the use of digital interventions that do not fit end user values and needs could affect intervention acceptance, adoption, and diffusion, indicating that a person-centered and iterative approach is needed [29,31,32,36].

A recent scoping review examining human-centered eHealth development indicated that because of the complexity of eHealth development, multiple strategies and methods should be combined in line with the research objectives when conducting such studies [115]. As such, a combination of theory, research-based, creative, and innovative methods, guided by the Double Diamond (ie, design thinking process) [65,66] and the CeHRes Roadmap [29,36] (Figure 2), can be applied to translate values and needs from ideation to the operationalization

of design features. The application of co-design and Agile development methods during this process may increase the chance that the values and needs of (future) users are met [43,45,110]. The tailoring and personalization of digital technologies in line with the values of individuals might be a promising motivational strategy for continued health behavior change [32]. Therefore, the human-centered and value-driven approach can be useful in strengthening the self-regulatory capacity and autonomous motivation required to achieve sustainable behavior change [26,32,120]. The application of theory-based approaches [52,55] when developing complex digital behavior change interventions such as eCHANGE can also facilitate specification and transparency of the internal structure of the technology and may enable the development and evaluation of high-quality and effective digital behavior change interventions [29,35,121].

Recommendations for Future Research

Future research should aim to investigate how digital technologies can be effective in facilitating sustainable behavior change and successful weight gain prevention and whether the identified design features, PSD principles, and BCTs contribute to long-term behavior change and sustainable health effects. In addition, while sustained technology engagement is not necessarily needed for a digital intervention to be useful, long-term weight loss maintenance supported by digital tools requires some degree of use. The degree of use needed and the design features that best support engagement and motivation for continued health behavior change and long-term weight maintenance are yet unknown. Therefore, future research should aim to evaluate technology engagement related to individual differences and the design features or intervention components that contribute to the creation of engaging technologies [122,123]. Hence, to optimize the impact of digital health interventions through personalized design, future research should also aim to provide more knowledge on how to tailor and personalize digital weight loss maintenance interventions (eg, design preferences) to support identified values and needs for long-term weight maintenance [32].

Finally, studies testing the actual use of the technology (eg, through log data analysis) may also provide knowledge related to end user preferences and needs (eg, latent, contextual, and future needs) and help identify which features may work best for whom and why to maintain weight. As such, a feasibility pilot trial may be a cost-effective way of acquiring more knowledge related to feasibility, desirability, and preliminary efficacy to support weight loss maintenance [54]. A feasibility pilot trial can also provide insights into user experiences and the actual use of the different components of technology to optimize and tailor digital interventions. Future research should also aim to evaluate the (long-term) effectiveness of weight loss maintenance and health-related outcomes through randomized controlled trials. To find the most effective combination of design features to support behavior change and weight loss maintenance over time, experimental designs, such as fractional factorial designs, could also be performed to test which features or combinations of features might work best for whom and when [29].

Strengths and Limitations

This study had some limitations. First, a few participating end users were still aiming to reduce their weight, although their initial goal for weight loss had been achieved. This could be a limitation for this study as end user needs during weight loss may differ from needs when aiming for weight maintenance [32,35]. However, the balance between weight loss and weight regain likely reflects real-life aspects and challenges related to weight loss maintenance, and this balance was also reflected in the technology design (eg, availability of *back-on-track* features).

Second, with respect to end user participation, more women than men participated in the study. There may be gender differences in weight loss maintenance strategies; therefore, this gender imbalance may limit the generalizability of the intervention to men, although several other key stakeholders participating were men, which might strengthen a more representative perspective.

Third, the participants in this study reported having a high school education or higher, which could indicate that this was a fairly well-educated sample. As such, the sample may not fully represent the diversity of the future end user population (ie, people with obesity aiming for weight maintenance after weight loss). However, existing research demonstrates inconsistent patterns of engagement and nonuse/attrition with respect to user characteristics (eg, education, age, and gender) [124,125]. In addition, the reported expectations and perceptions of a certain technology might not predict its actual use in practice [29]. To address these possible limitations, data from this study, as well as a previous study [32], were used to create user profiles (eg, need-based personas and user stories) representing potential end users and to guide the co-design and implementation process of adaptive, persuasive features to tailor the technology to individual preferences.

Fourth, because of limited time, resources, and privacy and security restrictions, some end user/stakeholder wishes (eg, personalized virtual coach, context-aware suggestions, social learning, and collaboration with other users) were not reflected

in the current MVP. On the basis of the nature of Agile development, the selected design features were prioritized with respect to practical (eg, privacy and security issues) and cost-effectiveness considerations (ie, value-based prioritization based on stakeholder input).

This study has numerous strengths. The focus on high multidisciplinary stakeholder involvement and end user values and needs are such strengths. Formative design results from co-design workshops, prototype validation, and usability testing with end users and other key stakeholders provided input with respect to what to deliver, as well as how, to support weight loss maintenance through design features (eg, content, function, and design). This highlights the importance of multidisciplinary collaboration between end users and other key stakeholders during the design and development of digital interventions to meet end user values and needs. Another strength is the focus on, and novelty of, how PSD principles and BCTs can be combined and implemented into design features of digital weight loss maintenance interventions. The application of the design features identified in this study could also be relevant for other areas requiring continued health behavior change, as the identified needs might reflect universal values or drivers for sustained behaviors [32].

Conclusions

To the best of our knowledge, this study is the first to combine PSD principles and BCTs into design features during the design and development of an evidence-informed digital behavior change intervention to support end user values and needs for long-term weight loss maintenance. The findings indicate that personalized digital weight maintenance interventions should aim to support health and well-being by including design features and strategies supporting the self-regulation of behaviors, thoughts and emotions, habit formation, autonomous motivation, competence, skills, and tailored support. The design and development of the eCHANGE intervention can provide valuable input for future design and tailoring of evidence-informed digital interventions, even beyond digital interventions in support of health behavior change and long-term weight loss maintenance.

Acknowledgments

This study was funded by the Department of Digital Health, Oslo University Hospital (OUH), and Vestfold Hospital Trust (VHT), Norway.

This study was part of the larger eCHANGE research project (principal investigator LSN), led by the Department of Digital Health Research at OUH, in collaboration with the Department of Endocrinology, Morbid Obesity, and Preventive Medicine at OUH; the Morbid Obesity Center at VHT; the Department of Physical Medicine and Rehabilitation at Sørlandet Hospital; and the Norwegian Association for People with Obesity, all located in Norway. Collaborators from municipalities surrounding the larger hospitals, as well as international researchers from the University of Twente, the Netherlands, and the Mayo Clinic, the United States, were also part of this project. The authors would especially like to thank all study participants, the research and development team, and the collaborating partners for their valuable contributions and participation in the research and development of the eCHANGE intervention. The authors would also like to thank the participating hospitals in Norway, VHT, OUH, and Sørlandet Hospital, for their support, effort, and interest in recruiting participants for this study.

Authors' Contributions

RAA was involved in all phases of the study and worked in close collaboration with LSN (principal investigator who led the intervention development) and the MO (product owner) during the design and development process. MLS, JH, and MMC participated in expert validation, consultation, and evidence-informed content development. LSN, JW, and JEWCVG-P provided feedback related to the analysis, results, and discussion. All the authors participated in the critical evaluation of the manuscript during several rounds before providing their final approval for submission.

Conflicts of Interest

None declared.

Multimedia Appendix 1

A design thinking process for eHealth design and development.

[\[PDF File \(Adobe PDF File\), 448 KB - humanfactors_v9i2e37372_app1.pdf\]](#)

Multimedia Appendix 2

Description of the need-based personas.

[\[PDF File \(Adobe PDF File\), 225 KB - humanfactors_v9i2e37372_app2.pdf\]](#)

Multimedia Appendix 3

Formative evaluation: detailed overview of methods and participants.

[\[PDF File \(Adobe PDF File\), 114 KB - humanfactors_v9i2e37372_app3.pdf\]](#)

Multimedia Appendix 4

Overview of end user demographics.

[\[PDF File \(Adobe PDF File\), 208 KB - humanfactors_v9i2e37372_app4.pdf\]](#)

Multimedia Appendix 5

Virtual Coach and Smart, Tailored Feedback system.

[\[PDF File \(Adobe PDF File\), 302 KB - humanfactors_v9i2e37372_app5.pdf\]](#)

Multimedia Appendix 6

Overview of design features, persuasive system design principles, and behavior change techniques combined and implemented.

[\[PDF File \(Adobe PDF File\), 222 KB - humanfactors_v9i2e37372_app6.pdf\]](#)

Multimedia Appendix 7

System Usability check.

[\[PDF File \(Adobe PDF File\), 128 KB - humanfactors_v9i2e37372_app7.pdf\]](#)

References

1. Middleton KR, Anton SD, Perri MG. Long-term adherence to health behavior change. *Am J Lifestyle Med* 2013 Jun 14;7(6):395-404 [FREE Full text] [doi: [10.1177/1559827613488867](https://doi.org/10.1177/1559827613488867)] [Medline: [27547170](https://pubmed.ncbi.nlm.nih.gov/27547170/)]
2. Lee I, Shiroma EJ, Lobelo F, Puska P, Blair SN, Katzmarzyk PT. Effect of physical inactivity on major non-communicable diseases worldwide: an analysis of burden of disease and life expectancy. *Lancet* 2012 Jul;380(9838):219-229. [doi: [10.1016/s0140-6736\(12\)61031-9](https://doi.org/10.1016/s0140-6736(12)61031-9)]
3. GBD 2017 Diet Collaborators. Health effects of dietary risks in 195 countries, 1990-2017: a systematic analysis for the Global Burden of Disease Study 2017. *Lancet* 2019 May 11;393(10184):1958-1972 [FREE Full text] [doi: [10.1016/S0140-6736\(19\)30041-8](https://doi.org/10.1016/S0140-6736(19)30041-8)] [Medline: [30954305](https://pubmed.ncbi.nlm.nih.gov/30954305/)]
4. GBD 2015 Obesity Collaborators, Afshin A, Forouzanfar MH, Reitsma MB, Sur P, Estep K, et al. Health Effects of Overweight and Obesity in 195 Countries over 25 Years. *N Engl J Med* 2017 Jul 06;377(1):13-27 [FREE Full text] [doi: [10.1056/NEJMoa1614362](https://doi.org/10.1056/NEJMoa1614362)] [Medline: [28604169](https://pubmed.ncbi.nlm.nih.gov/28604169/)]
5. Obesity and overweight. World Health Organization. URL: <https://www.who.int/news-room/fact-sheets/detail/obesity-and-overweight> [accessed 2022-05-03]
6. Roth G, Mensah G, Johnson C, Addolorato G, Ammirati E, Baddour L, GBD-NHLBI-JACC Global Burden of Cardiovascular Diseases Writing Group. Global burden of cardiovascular diseases and risk factors, 1990-2019: update from the GBD 2019 study. *J Am Coll Cardiol* 2020 Dec 22;76(25):2982-3021 [FREE Full text] [doi: [10.1016/j.jacc.2020.11.010](https://doi.org/10.1016/j.jacc.2020.11.010)] [Medline: [33309175](https://pubmed.ncbi.nlm.nih.gov/33309175/)]

7. Saeedi P, Petersohn I, Salpea P, Malanda B, Karuranga S, Unwin N, IDF Diabetes Atlas Committee. Global and regional diabetes prevalence estimates for 2019 and projections for 2030 and 2045: results from the international diabetes federation diabetes atlas, 9 edition. *Diabetes Res Clin Pract* 2019 Nov;157:107843. [doi: [10.1016/j.diabres.2019.107843](https://doi.org/10.1016/j.diabres.2019.107843)] [Medline: [31518657](https://pubmed.ncbi.nlm.nih.gov/31518657/)]
8. Wing R, Phelan S. Long-term weight loss maintenance. *Am J Clin Nutr* 2005 Jul;82(1 Suppl):222S-225S. [doi: [10.1093/ajcn/82.1.222S](https://doi.org/10.1093/ajcn/82.1.222S)] [Medline: [16002825](https://pubmed.ncbi.nlm.nih.gov/16002825/)]
9. Montesi L, El Ghoch M, Brodosi L, Calugi S, Marchesini G, Dalle Grave R. Long-term weight loss maintenance for obesity: a multidisciplinary approach. *Diabetes Metabolic Syndrome Obesity Targets Ther* 2016 Feb;9:37. [doi: [10.2147/dmso.s89836](https://doi.org/10.2147/dmso.s89836)]
10. MacLean PS, Wing RR, Davidson T, Epstein L, Goodpaster B, Hall KD, et al. NIH working group report: innovative research to improve maintenance of weight loss. *Obesity (Silver Spring)* 2015 Jan 02;23(1):7-15 [FREE Full text] [doi: [10.1002/oby.20967](https://doi.org/10.1002/oby.20967)] [Medline: [25469998](https://pubmed.ncbi.nlm.nih.gov/25469998/)]
11. Rosenbaum M, Leibel RL. Adaptive thermogenesis in humans. *Int J Obes (Lond)* 2010 Oct 11;34 Suppl 1(S1):S47-S55 [FREE Full text] [doi: [10.1038/ijo.2010.184](https://doi.org/10.1038/ijo.2010.184)] [Medline: [20935667](https://pubmed.ncbi.nlm.nih.gov/20935667/)]
12. Paixão C, Dias CM, Jorge R, Carraça EV, Yannakoulia M, de Zwaan M, et al. Successful weight loss maintenance: a systematic review of weight control registries. *Obes Rev* 2020 May;21(5):e13003. [doi: [10.1111/obr.13003](https://doi.org/10.1111/obr.13003)] [Medline: [32048787](https://pubmed.ncbi.nlm.nih.gov/32048787/)]
13. Wang YC, McPherson K, Marsh T, Gortmaker SL, Brown M. Health and economic burden of the projected obesity trends in the USA and the UK. *The Lancet* 2011 Aug;378(9793):815-825. [doi: [10.1016/s0140-6736\(11\)60814-3](https://doi.org/10.1016/s0140-6736(11)60814-3)]
14. Kastanias P, Mackenzie K, Robinson S, Wang W. Medical complications resulting from severe obesity. In: *Psychiatric Care in Severe Obesity*. Dordrecht/Heidelberg/New York/London: Springer; 2017.
15. Samdal GB, Eide GE, Barth T, Williams G, Meland E. Effective behaviour change techniques for physical activity and healthy eating in overweight and obese adults; systematic review and meta-regression analyses. *Int J Behav Nutr Phys Act* 2017 Mar 28;14(1):42 [FREE Full text] [doi: [10.1186/s12966-017-0494-y](https://doi.org/10.1186/s12966-017-0494-y)] [Medline: [28351367](https://pubmed.ncbi.nlm.nih.gov/28351367/)]
16. Gorin AA, Powers TA, Koestner R, Wing RR, Raynor HA. Autonomy support, self-regulation, and weight loss. *Health Psychol* 2014 Apr;33(4):332-339 [FREE Full text] [doi: [10.1037/a0032586](https://doi.org/10.1037/a0032586)] [Medline: [23730718](https://pubmed.ncbi.nlm.nih.gov/23730718/)]
17. Hall KD, Kahan S. Maintenance of lost weight and long-term management of obesity. *Med Clin North Am* 2018 Jan;102(1):183-197 [FREE Full text] [doi: [10.1016/j.mcna.2017.08.012](https://doi.org/10.1016/j.mcna.2017.08.012)] [Medline: [29156185](https://pubmed.ncbi.nlm.nih.gov/29156185/)]
18. Lawlor ER, Hughes CA, Duschinsky R, Pountain GD, Hill AJ, Griffin SJ, et al. Cognitive and behavioural strategies employed to overcome "lapses" and prevent "relapse" among weight-loss maintainers and regainers: a qualitative study. *Clin Obes* 2020 Oct 07;10(5):e12395 [FREE Full text] [doi: [10.1111/cob.12395](https://doi.org/10.1111/cob.12395)] [Medline: [32767708](https://pubmed.ncbi.nlm.nih.gov/32767708/)]
19. Phelan S, Halfman T, Pinto AM, Foster GD. Behavioral and psychological strategies of long-term weight loss maintainers in a widely available weight management program. *Obesity (Silver Spring)* 2020 Feb 23;28(2):421-428. [doi: [10.1002/oby.22685](https://doi.org/10.1002/oby.22685)] [Medline: [31970912](https://pubmed.ncbi.nlm.nih.gov/31970912/)]
20. Alamuddin N, Bakizada Z, Wadden TA. Management of obesity. *J Clin Oncol* 2016 Dec 10;34(35):4295-4305. [doi: [10.1200/jco.2016.66.8806](https://doi.org/10.1200/jco.2016.66.8806)]
21. Wing R, Espeland M, Clark J, Hazuda H, Knowler W, Pownall H, Action for Health in Diabetes (Look AHEAD) Study Group. Association of weight loss maintenance and weight regain on 4-year changes in CVD risk factors: the action for health in diabetes (look AHEAD) clinical trial. *Diabetes Care* 2016 Aug;39(8):1345-1355 [FREE Full text] [doi: [10.2337/dc16-0509](https://doi.org/10.2337/dc16-0509)] [Medline: [27271190](https://pubmed.ncbi.nlm.nih.gov/27271190/)]
22. Christiansen T, Bruun JM, Madsen EL, Richelsen B. Weight loss maintenance in severely obese adults after an intensive lifestyle intervention: 2- to 4-year follow-up. *Obesity (Silver Spring)* 2007 Feb;15(2):413-420 [FREE Full text] [doi: [10.1038/oby.2007.530](https://doi.org/10.1038/oby.2007.530)] [Medline: [17299115](https://pubmed.ncbi.nlm.nih.gov/17299115/)]
23. Greaves C, Poltawski L, Garside R, Briscoe S. Understanding the challenge of weight loss maintenance: a systematic review and synthesis of qualitative research on weight loss maintenance. *Health Psychol Rev* 2017 Jun 07;11(2):145-163. [doi: [10.1080/17437199.2017.1299583](https://doi.org/10.1080/17437199.2017.1299583)] [Medline: [28281891](https://pubmed.ncbi.nlm.nih.gov/28281891/)]
24. Pedersen S, Sniehotta FF, Sainsbury K, Evans EH, Marques MM, Stubbs RJ, et al. The complexity of self-regulating food intake in weight loss maintenance. A qualitative study among short- and long-term weight loss maintainers. *Soc Sci Med* 2018 Jul;208:18-24. [doi: [10.1016/j.socscimed.2018.05.016](https://doi.org/10.1016/j.socscimed.2018.05.016)] [Medline: [29753930](https://pubmed.ncbi.nlm.nih.gov/29753930/)]
25. Teixeira PJ, Carraça EV, Marques MM, Rutter H, Oppert J, De Bourdeaudhuij I, et al. Successful behavior change in obesity interventions in adults: a systematic review of self-regulation mediators. *BMC Med* 2015 Apr 16;13(1):84 [FREE Full text] [doi: [10.1186/s12916-015-0323-6](https://doi.org/10.1186/s12916-015-0323-6)] [Medline: [25907778](https://pubmed.ncbi.nlm.nih.gov/25907778/)]
26. Kwasnicka D, Dombrowski SU, White M, Sniehotta F. Theoretical explanations for maintenance of behaviour change: a systematic review of behaviour theories. *Health Psychol Rev* 2016 Sep 07;10(3):277-296 [FREE Full text] [doi: [10.1080/17437199.2016.1151372](https://doi.org/10.1080/17437199.2016.1151372)] [Medline: [26854092](https://pubmed.ncbi.nlm.nih.gov/26854092/)]
27. Okorodudu DE, Bosworth HB, Corsino L. Innovative interventions to promote behavioral change in overweight or obese individuals: a review of the literature. *Ann Med* 2015 May 10;47(3):179-185 [FREE Full text] [doi: [10.3109/07853890.2014.931102](https://doi.org/10.3109/07853890.2014.931102)] [Medline: [25011006](https://pubmed.ncbi.nlm.nih.gov/25011006/)]

28. Lentferink AJ, Oldenhuis HK, de Groot M, Polstra L, Velthuis H, van Gemert-Pijnen JE. Key components in eHealth interventions combining self-tracking and persuasive eCoaching to promote a healthier lifestyle: a scoping review. *J Med Internet Res* 2017 Aug 01;19(8):e277 [FREE Full text] [doi: [10.2196/jmir.7288](https://doi.org/10.2196/jmir.7288)] [Medline: [28765103](https://pubmed.ncbi.nlm.nih.gov/28765103/)]
29. van Gemert-Pijnen L, Kelders S, Kip H, Sanderman R, editors. *eHealth Research, Theory and Development A Multidisciplinary Approach*. London: Routledge; 2018.
30. Thomas Craig KJ, Morgan L, Chen C, Michie S, Fusco N, Snowdon J, et al. Systematic review of context-aware digital behavior change interventions to improve health. *Transl Behav Med* 2021 May 25;11(5):1037-1048 [FREE Full text] [doi: [10.1093/tbm/ibaa099](https://doi.org/10.1093/tbm/ibaa099)] [Medline: [33085767](https://pubmed.ncbi.nlm.nih.gov/33085767/)]
31. Michie S, Yardley L, West R, Patrick K, Greaves F. Developing and evaluating digital interventions to promote behavior change in health and health care: recommendations resulting from an international workshop. *J Med Internet Res* 2017 Jun 29;19(6):e232 [FREE Full text] [doi: [10.2196/jmir.7126](https://doi.org/10.2196/jmir.7126)] [Medline: [28663162](https://pubmed.ncbi.nlm.nih.gov/28663162/)]
32. Asbjørnsen RA, Wentzel J, Smedsrød ML, Hjelmæsæth J, Clark MM, Solberg Nes L, et al. Identifying persuasive design principles and behavior change techniques supporting end user values and needs in eHealth interventions for long-term weight loss maintenance: qualitative study. *J Med Internet Res* 2020 Nov 30;22(11):e22598 [FREE Full text] [doi: [10.2196/22598](https://doi.org/10.2196/22598)] [Medline: [33252347](https://pubmed.ncbi.nlm.nih.gov/33252347/)]
33. Kwasnicka D, Dombrowski SU, White M, Sniehotta FF. 'It's not a diet, it's a lifestyle': a longitudinal, data-prompted interview study of weight loss maintenance. *Psychol Health* 2019 Aug 23;34(8):963-982. [doi: [10.1080/08870446.2019.1579913](https://doi.org/10.1080/08870446.2019.1579913)] [Medline: [30905184](https://pubmed.ncbi.nlm.nih.gov/30905184/)]
34. Hutchesson MJ, Rollo ME, Krukowski R, Ells L, Harvey J, Morgan PJ, et al. eHealth interventions for the prevention and treatment of overweight and obesity in adults: a systematic review with meta-analysis. *Obes Rev* 2015 May 05;16(5):376-392. [doi: [10.1111/obr.12268](https://doi.org/10.1111/obr.12268)] [Medline: [25753009](https://pubmed.ncbi.nlm.nih.gov/25753009/)]
35. Asbjørnsen RA, Smedsrød ML, Solberg Nes L, Wentzel J, Varsi C, Hjelmæsæth J, et al. Persuasive system design principles and behavior change techniques to stimulate motivation and adherence in electronic health interventions to support weight loss maintenance: scoping review. *J Med Internet Res* 2019 Jun 21;21(6):e14265 [FREE Full text] [doi: [10.2196/14265](https://doi.org/10.2196/14265)] [Medline: [31228174](https://pubmed.ncbi.nlm.nih.gov/31228174/)]
36. van Gemert-Pijnen JE, Nijland N, van Limburg M, Ossebaard HC, Kelders SM, Eysenbach G, et al. A holistic framework to improve the uptake and impact of eHealth technologies. *J Med Internet Res* 2011 Dec 05;13(4):e111 [FREE Full text] [doi: [10.2196/jmir.1672](https://doi.org/10.2196/jmir.1672)] [Medline: [22155738](https://pubmed.ncbi.nlm.nih.gov/22155738/)]
37. Yardley L, Morrison L, Bradbury K, Muller I. The person-based approach to intervention development: application to digital health-related behavior change interventions. *J Med Internet Res* 2015 Jan 30;17(1):e30 [FREE Full text] [doi: [10.2196/jmir.4055](https://doi.org/10.2196/jmir.4055)] [Medline: [25639757](https://pubmed.ncbi.nlm.nih.gov/25639757/)]
38. Simonsen J, Robertson T, editors. *Routledge International Handbook of Participatory Design*. London: Routledge; 2013.
39. Kok G, Peters LW, Ruiters RA. Planning theory- and evidence-based behavior change interventions: a conceptual review of the intervention mapping protocol. *Psicol Reflex Crit* 2017 Oct 18;30(1):19 [FREE Full text] [doi: [10.1186/s41155-017-0072-x](https://doi.org/10.1186/s41155-017-0072-x)] [Medline: [32026109](https://pubmed.ncbi.nlm.nih.gov/32026109/)]
40. Altman M, Huang TT, Breland JY. Design thinking in health care. *Prev Chronic Dis* 2018 Sep 27;15:E117 [FREE Full text] [doi: [10.5888/pcd15.180128](https://doi.org/10.5888/pcd15.180128)] [Medline: [30264690](https://pubmed.ncbi.nlm.nih.gov/30264690/)]
41. Stickdorn M, Schneider J. *This is Service Design Thinking: Basics, Tools, Cases*. Hoboken, New Jersey, United States: Wiley; 2012.
42. Casanova P. Agile processes: a unifying approach for the future of projects. In: *Proceedings of the PMI® Global Congress 2013*. 2013 Presented at: PMI® Global Congress 2013; Apr 24, 2013; Istanbul, Turkey.
43. Hekler EB, Klasnja P, Riley WT, Buman MP, Huberty J, Rivera DE, et al. Agile science: creating useful products for behavior change in the real world. *Transl Behav Med* 2016 Jun 26;6(2):317-328 [FREE Full text] [doi: [10.1007/s13142-016-0395-7](https://doi.org/10.1007/s13142-016-0395-7)] [Medline: [27357001](https://pubmed.ncbi.nlm.nih.gov/27357001/)]
44. Patrick K, Hekler EB, Estrin D, Mohr DC, Riper H, Crane D, et al. The pace of technologic change: implications for digital health behavior intervention research. *Am J Preventive Med* 2016 Nov;51(5):816-824. [doi: [10.1016/j.amepre.2016.05.001](https://doi.org/10.1016/j.amepre.2016.05.001)]
45. Sanders EB, Stappers PJ. Co-creation and the new landscapes of design. *CoDesign* 2008 Mar;4(1):5-18. [doi: [10.1080/15710880701875068](https://doi.org/10.1080/15710880701875068)]
46. Roberts JP, Fisher TR, Trowbridge MJ, Bent C. A design thinking framework for healthcare management and innovation. *Healthc (Amst)* 2016 Mar;4(1):11-14. [doi: [10.1016/j.hjdsi.2015.12.002](https://doi.org/10.1016/j.hjdsi.2015.12.002)] [Medline: [27001093](https://pubmed.ncbi.nlm.nih.gov/27001093/)]
47. Ries E. The lean startup: how today's entrepreneurs use continuous innovation to create radically successful businesses. *J Prod Innov Manag* 2012 Mar 19;29(3):508-509. [doi: [10.1111/j.1540-5885.2012.00920_2.x](https://doi.org/10.1111/j.1540-5885.2012.00920_2.x)]
48. Osterwalder A, Pigneur Y, Bernarda G, Smith A. *Value Proposition Design: How to Create Products and Services Customers Want*. Hoboken, New Jersey, United States: Wiley; 2014.
49. Friedman B, Kahn P, Borning A, Hultgren A. Value sensitive design and information systems. In: *Early Engagement and New Technologies: Opening Up the Laboratory*. Dordrecht/Heidelberg/New York/London: Springer; 2013.
50. Van Velsen L, Wentzel J, Van Gemert-Pijnen JE. Designing eHealth that matters via a multidisciplinary requirements development approach. *JMIR Res Protoc* 2013 Jun 24;2(1):e21 [FREE Full text] [doi: [10.2196/resprot.2547](https://doi.org/10.2196/resprot.2547)] [Medline: [23796508](https://pubmed.ncbi.nlm.nih.gov/23796508/)]

51. Michie S, Abraham C, Eccles MP, Francis JJ, Hardeman W, Johnston M. Strengthening evaluation and implementation by specifying components of behaviour change interventions: a study protocol. *Implement Sci* 2011 Feb 07;6(1):10 [FREE Full text] [doi: [10.1186/1748-5908-6-10](https://doi.org/10.1186/1748-5908-6-10)] [Medline: [21299860](https://pubmed.ncbi.nlm.nih.gov/21299860/)]
52. Michie S, Richardson M, Johnston M, Abraham C, Francis J, Hardeman W, et al. The behavior change technique taxonomy (v1) of 93 hierarchically clustered techniques: building an international consensus for the reporting of behavior change interventions. *Ann Behav Med* 2013 Aug 20;46(1):81-95. [doi: [10.1007/s12160-013-9486-6](https://doi.org/10.1007/s12160-013-9486-6)] [Medline: [23512568](https://pubmed.ncbi.nlm.nih.gov/23512568/)]
53. Dombrowski SU, Sniehotta FF, Avenell A, Johnston M, MacLennan G, Araújo-Soares V. Identifying active ingredients in complex behavioural interventions for obese adults with obesity-related co-morbidities or additional risk factors for co-morbidities: a systematic review. *Health Psychol Rev* 2012 Mar;6(1):7-32. [doi: [10.1080/17437199.2010.513298](https://doi.org/10.1080/17437199.2010.513298)]
54. Skivington K, Matthews L, Simpson SA, Craig P, Baird J, Blazeby JM, et al. A new framework for developing and evaluating complex interventions: update of medical research council guidance. *BMJ* 2021 Sep 30;374:n2061 [FREE Full text] [doi: [10.1136/bmj.n2061](https://doi.org/10.1136/bmj.n2061)] [Medline: [34593508](https://pubmed.ncbi.nlm.nih.gov/34593508/)]
55. Oinas-Kukkonen H, Harjumaa M. Persuasive systems design: key issues, process model, and system features. *Commun Assoc Inf Syst* 2009;24:485-500. [doi: [10.17705/1cais.02428](https://doi.org/10.17705/1cais.02428)]
56. Kelders SM, van Gemert-Pijnen JE, Werkman A, Seydel ER. Evaluation of a web-based lifestyle coach designed to maintain a healthy bodyweight. *J Telemed Telecare* 2010 Jan 19;16(1):3-7. [doi: [10.1258/jtt.2009.001003](https://doi.org/10.1258/jtt.2009.001003)] [Medline: [20086259](https://pubmed.ncbi.nlm.nih.gov/20086259/)]
57. Geuens J, Geurts L, Gerling K, Croon R, Abeele V. A dyad of lenses for the motivational design of mHealth: bridging the gap between health theory and app design. In: Proceedings of the 2019 IEEE International Conference on Healthcare Informatics (ICHI). 2019 Presented at: 2019 IEEE International Conference on Healthcare Informatics (ICHI); Jun 10-13, 2019; Xi'an, China URL: <https://ieeexplore.ieee.org/document/8904839> [doi: [10.1109/ichi.2019.8904839](https://doi.org/10.1109/ichi.2019.8904839)]
58. Sittig S, McGowan A, Iyengar S. Extensive review of persuasive system design categories and principles: behavioral obesity interventions. *J Med Syst* 2020 Jun 05;44(7):128. [doi: [10.1007/s10916-020-01591-w](https://doi.org/10.1007/s10916-020-01591-w)] [Medline: [32500161](https://pubmed.ncbi.nlm.nih.gov/32500161/)]
59. Cruz-Martínez RR, Wentzel J, Asbjørnsen RA, Noort PD, van Niekerk JM, Sanderman R, et al. Supporting self-management of cardiovascular diseases through remote monitoring technologies: metaethnography review of frameworks, models, and theories used in research and development. *J Med Internet Res* 2020 May 21;22(5):e16157 [FREE Full text] [doi: [10.2196/16157](https://doi.org/10.2196/16157)] [Medline: [32436852](https://pubmed.ncbi.nlm.nih.gov/32436852/)]
60. Taj F, Klein MC, van Halteren A. Digital health behavior change technology: bibliometric and scoping review of two decades of research. *JMIR Mhealth Uhealth* 2019 Dec 13;7(12):e13311 [FREE Full text] [doi: [10.2196/13311](https://doi.org/10.2196/13311)] [Medline: [31833836](https://pubmed.ncbi.nlm.nih.gov/31833836/)]
61. Moller AC, Merchant G, Conroy DE, West R, Hekler E, Kugler KC, et al. Applying and advancing behavior change theories and techniques in the context of a digital health revolution: proposals for more effectively realizing untapped potential. *J Behav Med* 2017 Feb 5;40(1):85-98 [FREE Full text] [doi: [10.1007/s10865-016-9818-7](https://doi.org/10.1007/s10865-016-9818-7)] [Medline: [28058516](https://pubmed.ncbi.nlm.nih.gov/28058516/)]
62. Sporrel K, Nibbeling N, Wang S, Etema D, Simons M. Unraveling mobile health exercise interventions for adults: scoping review on the implementations and designs of persuasive strategies. *JMIR Mhealth Uhealth* 2021 Jan 18;9(1):e16282 [FREE Full text] [doi: [10.2196/16282](https://doi.org/10.2196/16282)] [Medline: [33459598](https://pubmed.ncbi.nlm.nih.gov/33459598/)]
63. Morrison LG, Yardley L, Powell J, Michie S. What design features are used in effective e-health interventions? A review using techniques from critical interpretive synthesis. *Telemed J E Health* 2012 Mar;18(2):137-144. [doi: [10.1089/tmj.2011.0062](https://doi.org/10.1089/tmj.2011.0062)] [Medline: [22381060](https://pubmed.ncbi.nlm.nih.gov/22381060/)]
64. Schoeppe S, Alley S, Van Lippevelde W, Bray NA, Williams SL, Duncan MJ, et al. Efficacy of interventions that use apps to improve diet, physical activity and sedentary behaviour: a systematic review. *Int J Behav Nutr Phys Act* 2016 Dec 07;13(1):127 [FREE Full text] [doi: [10.1186/s12966-016-0454-y](https://doi.org/10.1186/s12966-016-0454-y)] [Medline: [27927218](https://pubmed.ncbi.nlm.nih.gov/27927218/)]
65. What is the framework for innovation? Design Council's evolved Double Diamond. Design Council. URL: <https://www.designcouncil.org.uk/news-opinion/what-framework-innovation-design-councils-evolved-double-diamond> [accessed 2022-05-03]
66. Council B. The Double Diamond. Design Council. URL: <https://www.designcouncil.org.uk/news-opinion/double-diamond-universally-accepted-depiction-design-process> [accessed 2022-05-10]
67. What are overweight and obesity. World Health Organization. URL: <https://www.who.int/news-room/fact-sheets/detail/obesity-and-overweight> [accessed 2022-05-10]
68. Polaine A, Reason B, Løvlie L. Service Design: From Insight to Implementation. Brooklyn, New York, United States: Rosenfeld Media; 2013.
69. LeRouge C, Ma J, Sneha S, Tolle K. User profiles and personas in the design and development of consumer health technologies. *Int J Med Inform* 2013 Nov;82(11):e251-e268. [doi: [10.1016/j.ijmedinf.2011.03.006](https://doi.org/10.1016/j.ijmedinf.2011.03.006)] [Medline: [21481635](https://pubmed.ncbi.nlm.nih.gov/21481635/)]
70. Holden RJ, Kulanthaivel A, Purkayastha S, Goggins KM, Kripalani S. Know thy eHealth user: development of biopsychosocial personas from a study of older adults with heart failure. *Int J Med Inform* 2017 Dec;108:158-167 [FREE Full text] [doi: [10.1016/j.ijmedinf.2017.10.006](https://doi.org/10.1016/j.ijmedinf.2017.10.006)] [Medline: [29132622](https://pubmed.ncbi.nlm.nih.gov/29132622/)]
71. Varkevisser RD, van Stralen MM, Kroeze W, Ket JC, Steenhuis IH. Determinants of weight loss maintenance: a systematic review. *Obes Rev* 2019 Feb 16;20(2):171-211 [FREE Full text] [doi: [10.1111/obr.12772](https://doi.org/10.1111/obr.12772)] [Medline: [30324651](https://pubmed.ncbi.nlm.nih.gov/30324651/)]
72. Elfhag K, Rossner S. Who succeeds in maintaining weight loss? A conceptual review of factors associated with weight loss maintenance and weight regain. *Obesity Rev* 2005 Feb;6(1):67-85. [doi: [10.1111/j.1467-789x.2005.00170.x](https://doi.org/10.1111/j.1467-789x.2005.00170.x)]

73. Wing RR, Hill JO. Successful weight loss maintenance. *Annu Rev Nutr* 2001;21:323-341. [doi: [10.1146/annurev.nutr.21.1.323](https://doi.org/10.1146/annurev.nutr.21.1.323)] [Medline: [11375440](https://pubmed.ncbi.nlm.nih.gov/11375440/)]
74. Thomas JG, Bond DS, Phelan S, Hill JO, Wing RR. Weight-loss maintenance for 10 years in the national weight control registry. *Am J Prev Med* 2014 Jan;46(1):17-23. [doi: [10.1016/j.amepre.2013.08.019](https://doi.org/10.1016/j.amepre.2013.08.019)] [Medline: [24355667](https://pubmed.ncbi.nlm.nih.gov/24355667/)]
75. Anbefalinger om kosthold, ernæring og fysisk aktivitet. Helsedirektoratet. URL: <https://tinyurl.com/4drde7jb> [accessed 2022-05-12]
76. Helsedirektoratet. Forebygging, utredning og behandling av overvekt og fedme hos voksne: nasjonale retningslinjer for primærhelsetjenesten. Oslo, Norway: Helsedirektoratet; 2011.
77. The National Weight Control Registry homepage. The National Weight Control Registry. URL: <http://www.nwcr.ws/> [accessed 2022-05-04]
78. Look AHEAD Research Group, Wadden TA, West DS, Delahanty L, Jakicic J, Rejeski J, et al. The Look AHEAD study: a description of the lifestyle intervention and the evidence supporting it. *Obesity (Silver Spring)* 2006 May;14(5):737-752 [FREE Full text] [doi: [10.1038/oby.2006.84](https://doi.org/10.1038/oby.2006.84)] [Medline: [16855180](https://pubmed.ncbi.nlm.nih.gov/16855180/)]
79. Jensen MD, Ryan DH, Donato KA, Apovian CM, Ard JD, Comuzzie AG, et al. Executive summary: guidelines (2013) for the management of overweight and obesity in adults. *Obesity* 2014 Jun 24;22(S2):S5-39. [doi: [10.1002/oby.20821](https://doi.org/10.1002/oby.20821)]
80. Jensen M, Ryan D, Apovian C, Ard J, Comuzzie A, Donato K, American College of Cardiology/American Heart Association Task Force on Practice Guidelines, Obesity Society. 2013 AHA/ACC/TOS guideline for the management of overweight and obesity in adults: a report of the American College of Cardiology/American Heart Association Task Force on Practice Guidelines and The Obesity Society. *Circulation* 2014 Jun 24;129(25 Suppl 2):S102-S138 [FREE Full text] [doi: [10.1161/01.cir.0000437739.71477.ee](https://doi.org/10.1161/01.cir.0000437739.71477.ee)] [Medline: [24222017](https://pubmed.ncbi.nlm.nih.gov/24222017/)]
81. Interpretation and overview of test procedures for WCAG 2.0 A and AA. uutilsynet. URL: <https://www.uutilsynet.no/english/interpretation-and-overview-test-procedures-wcag-20-and-aa/138> [accessed 2022-05-04]
82. Software development with data protection by design and by default. Datatilsynet. URL: <https://www.datatilsynet.no/en/about-privacy/virksomhetenes-plikter/innebygd-personvern/data-protection-by-design-and-by-default/?print=true> [accessed 2022-05-04]
83. OWASP application security verification standard. OWASP. URL: <https://owasp.org/www-project-application-security-verification-standard/> [accessed 2022-05-04]
84. Sauro J. A Practical Guide to the System Usability Scale: Background, Benchmarks & Best Practices. Scotts Valley, California, United States: Createspace Independent Pub; 2011.
85. Meso P, Jain R. Agile software development: adaptive systems principles and best practices. *Inf Sys Manag* 2006 Jun;23(3):19-30. [doi: [10.1201/1078.10580530/46108.23.3.20060601/93704.3](https://doi.org/10.1201/1078.10580530/46108.23.3.20060601/93704.3)]
86. Gale RC, Wu J, Erhardt T, Bounthavong M, Reardon CM, Damschroder LJ, et al. Comparison of rapid vs in-depth qualitative analytic methods from a process evaluation of academic detailing in the veterans health administration. *Implement Sci* 2019 Feb 01;14(1):11 [FREE Full text] [doi: [10.1186/s13012-019-0853-y](https://doi.org/10.1186/s13012-019-0853-y)] [Medline: [30709368](https://pubmed.ncbi.nlm.nih.gov/30709368/)]
87. Hsieh H, Shannon SE. Three approaches to qualitative content analysis. *Qual Health Res* 2005 Nov 01;15(9):1277-1288. [doi: [10.1177/1049732305276687](https://doi.org/10.1177/1049732305276687)] [Medline: [16204405](https://pubmed.ncbi.nlm.nih.gov/16204405/)]
88. Ledel Solem IK, Varsi C, Eide H, Kristjansdottir OB, Børøsund E, Schreurs KM, et al. A user-centered approach to an evidence-based electronic health pain management intervention for people with chronic pain: design and development of EPIO. *J Med Internet Res* 2020 Jan 21;22(1):e15889 [FREE Full text] [doi: [10.2196/15889](https://doi.org/10.2196/15889)] [Medline: [31961331](https://pubmed.ncbi.nlm.nih.gov/31961331/)]
89. Miller W, Rollnick S. Motivational Interviewing, Third Edition Helping People Change. New York, United States: Guilford Publications; 1991.
90. Wing RR, Tate DF, Gorin AA, Raynor HA, Fava JL. A self-regulation program for maintenance of weight loss. *N Engl J Med* 2006 Oct 12;355(15):1563-1571. [doi: [10.1056/nejmoa061883](https://doi.org/10.1056/nejmoa061883)]
91. Bandura A. Social cognitive theory of self-regulation. *Org Behav Human Decision Processes* 1991 Dec;50(2):248-287. [doi: [10.1016/0749-5978\(91\)90022-1](https://doi.org/10.1016/0749-5978(91)90022-1)]
92. Ryan RM, Deci EL. Self-determination theory and the facilitation of intrinsic motivation, social development, and well-being. *Am Psychol* 2000;55(1):68-78. [doi: [10.1037/0003-066x.55.1.68](https://doi.org/10.1037/0003-066x.55.1.68)]
93. Hayes S, Strosahl KD, Wilson KG. Acceptance and Commitment Therapy: an Experiential Approach to Behavior Change. New York, United States: Guilford Press; 1999.
94. Beck J. Cognitive Behavior Therapy Third Edition Basics and Beyond. New York, United States: Guilford Press; 2020.
95. Deci EL, Ryan RM. Intrinsic Motivation and Self-Determination in Human Behavior. Dordrecht/Heidelberg/New York/London: Springer; 1985.
96. Bandura A. Self-efficacy: The Exercise of Control. United States: W H Freeman/Times Books/Henry Holt & Co; 1997.
97. Hartmann-Boyce J, Nourse R, Boylan A, Jebb SA, Aveyard P. Experiences of reframing during self-directed weight loss and weight loss maintenance: systematic review of qualitative studies. *Appl Psychol Health Well Being* 2018 Jul 01;10(2):309-329 [FREE Full text] [doi: [10.1111/aphw.12132](https://doi.org/10.1111/aphw.12132)] [Medline: [29856139](https://pubmed.ncbi.nlm.nih.gov/29856139/)]
98. Spreckley M, Seidell J, Halberstadt J. Perspectives into the experience of successful, substantial long-term weight-loss maintenance: a systematic review. *Int J Qual Stud Health Well-being* 2021 Dec 17;16(1):1862481 [FREE Full text] [doi: [10.1080/17482631.2020.1862481](https://doi.org/10.1080/17482631.2020.1862481)] [Medline: [33455563](https://pubmed.ncbi.nlm.nih.gov/33455563/)]

99. Silva MN, Vieira PN, Coutinho SR, Minderico CS, Matos MG, Sardinha LB, et al. Using self-determination theory to promote physical activity and weight control: a randomized controlled trial in women. *J Behav Med* 2010 Apr 11;33(2):110-122. [doi: [10.1007/s10865-009-9239-y](https://doi.org/10.1007/s10865-009-9239-y)] [Medline: [20012179](https://pubmed.ncbi.nlm.nih.gov/20012179/)]
100. van der Weiden A, Benjamins J, Gillebaart M, Ybema JF, de Ridder D. How to form good habits? A longitudinal field study on the role of self-control in habit formation. *Front Psychol* 2020 Mar 27;11:560 [FREE Full text] [doi: [10.3389/fpsyg.2020.00560](https://doi.org/10.3389/fpsyg.2020.00560)] [Medline: [32292376](https://pubmed.ncbi.nlm.nih.gov/32292376/)]
101. Natvik E, Råheim M, Andersen JR, Moltu C. An experientially derived model of flexible and intentional actions for weight loss maintenance after severe obesity. *Front Psychol* 2019 Nov 13;10:2503 [FREE Full text] [doi: [10.3389/fpsyg.2019.02503](https://doi.org/10.3389/fpsyg.2019.02503)] [Medline: [31798491](https://pubmed.ncbi.nlm.nih.gov/31798491/)]
102. Dusseldorp E, van Genugten L, van Buuren S, Verheijden MW, van Empelen P. Combinations of techniques that effectively change health behavior: evidence from Meta-CART analysis. *Health Psychol* 2014 Dec;33(12):1530-1540. [doi: [10.1037/hea0000018](https://doi.org/10.1037/hea0000018)] [Medline: [24274802](https://pubmed.ncbi.nlm.nih.gov/24274802/)]
103. Haywood D, Lawrence BJ, Baughman FD, Mullan BA. A conceptual model of long-term weight loss maintenance: the importance of cognitive, empirical and computational approaches. *Int J Environ Res Public Health* 2021 Jan 13;18(2):635 [FREE Full text] [doi: [10.3390/ijerph18020635](https://doi.org/10.3390/ijerph18020635)] [Medline: [33451110](https://pubmed.ncbi.nlm.nih.gov/33451110/)]
104. Gardner B, Richards R, Lally P, Rebar A, Thwaite T, Beeken RJ. Breaking habits or breaking habitual behaviours? Old habits as a neglected factor in weight loss maintenance. *Appetite* 2021 Jul 01;162:105183. [doi: [10.1016/j.appet.2021.105183](https://doi.org/10.1016/j.appet.2021.105183)] [Medline: [33651994](https://pubmed.ncbi.nlm.nih.gov/33651994/)]
105. Roordink EM, Steenhuis IH, Kroeze W, Schoonmade LJ, Sniehotta FF, van Stralen MM. Predictors of lapse and relapse in physical activity and dietary behaviour: a systematic search and review on prospective studies. *Psychol Health* 2021 Dec 01:1-24. [doi: [10.1080/08870446.2021.1981900](https://doi.org/10.1080/08870446.2021.1981900)] [Medline: [34851220](https://pubmed.ncbi.nlm.nih.gov/34851220/)]
106. Cleo G, Glasziou P, Beller E, Isenring E, Thomas R. Habit-based interventions for weight loss maintenance in adults with overweight and obesity: a randomized controlled trial. *Int J Obes (Lond)* 2019 Feb 23;43(2):374-383. [doi: [10.1038/s41366-018-0067-4](https://doi.org/10.1038/s41366-018-0067-4)] [Medline: [29686382](https://pubmed.ncbi.nlm.nih.gov/29686382/)]
107. Cleo G, Isenring E, Thomas R, Glasziou P. Could habits hold the key to weight loss maintenance? A narrative review. *J Hum Nutr Diet* 2017 Oct 02;30(5):655-664. [doi: [10.1111/jhn.12456](https://doi.org/10.1111/jhn.12456)] [Medline: [28150402](https://pubmed.ncbi.nlm.nih.gov/28150402/)]
108. Awoke MA, Harrison CL, Martin J, Misso ML, Lim S, Moran LJ. Behaviour change techniques in weight gain prevention interventions in adults of reproductive age: meta-analysis and meta-regression. *Nutrients* 2022 Jan 03;14(1):209 [FREE Full text] [doi: [10.3390/nu14010209](https://doi.org/10.3390/nu14010209)] [Medline: [35011084](https://pubmed.ncbi.nlm.nih.gov/35011084/)]
109. Sheeran P, Webb TL. The intention-behavior gap. *Social Personality Psychol Compass* 2016 Sep 05;10(9):503-518. [doi: [10.1111/spc3.12265](https://doi.org/10.1111/spc3.12265)]
110. Petermans A, Cain R, editors. *Design for Wellbeing: An Applied Approach*. London: Routledge; 2019.
111. Olander EK, Fletcher H, Williams S, Atkinson L, Turner A, French DP. What are the most effective techniques in changing obese individuals' physical activity self-efficacy and behaviour: a systematic review and meta-analysis. *Int J Behav Nutr Phys Act* 2013 Mar 03;10(1):29 [FREE Full text] [doi: [10.1186/1479-5868-10-29](https://doi.org/10.1186/1479-5868-10-29)] [Medline: [23452345](https://pubmed.ncbi.nlm.nih.gov/23452345/)]
112. Carey R, Connell L, Johnston M, Rothman A, de Bruin M, Kelly M, et al. Behavior change techniques and their mechanisms of action: a synthesis of links described in published intervention literature. *Ann Behav Med* 2019 Jul 17;53(8):693-707 [FREE Full text] [doi: [10.1093/abm/kay078](https://doi.org/10.1093/abm/kay078)] [Medline: [30304386](https://pubmed.ncbi.nlm.nih.gov/30304386/)]
113. Michie S, West R, Campbell R, Brown J, Gainforth H. *ABC of Behaviour Change Theories*. United Kingdom: Silverback Publishing; 2014.
114. Desmet PA, Pohlmeier AE. Positive design: an introduction to design for subjective well-being. *Int J Design* 2013;7(3):5-19.
115. Kip H, Keizer J, da Silva MC, Beerlage-de Jong N, Köhler N, Kelders SM. Methods for human-centered eHealth development: narrative scoping review. *J Med Internet Res* 2022 Jan 27;24(1):e31858 [FREE Full text] [doi: [10.2196/31858](https://doi.org/10.2196/31858)] [Medline: [35084359](https://pubmed.ncbi.nlm.nih.gov/35084359/)]
116. Brindal E, Hendrie GA, Freyne J, Noakes M. A mobile phone app designed to support weight loss maintenance and well-being (MotiMate): randomized controlled trial. *JMIR Mhealth Uhealth* 2019 Sep 04;7(9):e12882 [FREE Full text] [doi: [10.2196/12882](https://doi.org/10.2196/12882)] [Medline: [31486407](https://pubmed.ncbi.nlm.nih.gov/31486407/)]
117. Evans EH, Araújo-Soares V, Adamson A, Batterham AM, Brown H, Campbell M, et al. The NULevel trial of a scalable, technology-assisted weight loss maintenance intervention for obese adults after clinically significant weight loss: study protocol for a randomised controlled trial. *Trials* 2015;16:421. [doi: [10.1186/s13063-015-0931-7](https://doi.org/10.1186/s13063-015-0931-7)]
118. Marques MM, Matos M, Mattila E, Encantado J, Duarte C, Teixeira PJ, et al. A theory- and evidence-based digital intervention tool for weight loss maintenance (NoHoW toolkit): systematic development and refinement study. *J Med Internet Res* 2021 Dec 03;23(12):e25305 [FREE Full text] [doi: [10.2196/25305](https://doi.org/10.2196/25305)] [Medline: [34870602](https://pubmed.ncbi.nlm.nih.gov/34870602/)]
119. Richards R, Jones RA, Whittle F, Hughes CA, Hill AJ, Lawlor ER, et al. Development of a web-based, guided self-help, acceptance and commitment therapy-based intervention for weight loss maintenance: evidence-, theory-, and person-based approach. *JMIR Form Res* 2022 Jan 07;6(1):e31801 [FREE Full text] [doi: [10.2196/31801](https://doi.org/10.2196/31801)] [Medline: [34994698](https://pubmed.ncbi.nlm.nih.gov/34994698/)]
120. Baumeister RF, Heatherton TF, Tice DM. *Losing Control How and Why People Fail at Self-Regulation*. Amsterdam, Netherlands: Elsevier Science; 1994.

121. Michie S, van Stralen MM, West R. The behaviour change wheel: a new method for characterising and designing behaviour change interventions. *Implement Sci* 2011 Apr 23;6(1):42 [FREE Full text] [doi: [10.1186/1748-5908-6-42](https://doi.org/10.1186/1748-5908-6-42)] [Medline: [21513547](https://pubmed.ncbi.nlm.nih.gov/21513547/)]
122. Kelders SM, Kip H, Greeff J. Psychometric evaluation of the TWente Engagement with Ehealth Technologies Scale (TWEETS): evaluation study. *J Med Internet Res* 2020 Oct 09;22(10):e17757 [FREE Full text] [doi: [10.2196/17757](https://doi.org/10.2196/17757)] [Medline: [33021487](https://pubmed.ncbi.nlm.nih.gov/33021487/)]
123. Yardley L, Spring BJ, Riper H, Morrison LG, Crane DH, Curtis K, et al. Understanding and promoting effective engagement with digital behavior change interventions. *Am J Prev Med* 2016 Nov;51(5):833-842. [doi: [10.1016/j.amepre.2016.06.015](https://doi.org/10.1016/j.amepre.2016.06.015)] [Medline: [27745683](https://pubmed.ncbi.nlm.nih.gov/27745683/)]
124. Neve MJ, Collins CE, Morgan PJ. Dropout, nonusage attrition, and pretreatment predictors of nonusage attrition in a commercial web-based weight loss program. *J Med Internet Res* 2010 Dec 14;12(4):e69 [FREE Full text] [doi: [10.2196/jmir.1640](https://doi.org/10.2196/jmir.1640)] [Medline: [21156470](https://pubmed.ncbi.nlm.nih.gov/21156470/)]
125. Sieverink F, Kelders S, Poel M, van Gemert-Pijnen L. Opening the black box of electronic health: collecting, analyzing, and interpreting log data. *JMIR Res Protoc* 2017 Aug 07;6(8):e156 [FREE Full text] [doi: [10.2196/resprot.6452](https://doi.org/10.2196/resprot.6452)] [Medline: [28784592](https://pubmed.ncbi.nlm.nih.gov/28784592/)]

Abbreviations

BCT: behavior change technique
CeHRes: Center for eHealth Research and Disease Management
MVP: minimal viable product
OUH: Oslo University Hospital
PI: principal investigator
PSD: persuasive system design
VHT: Vestfold Hospital Trust

Edited by A Kushniruk; submitted 17.02.22; peer-reviewed by M Canazei, L Guo; comments to author 15.03.22; revised version received 29.03.22; accepted 22.04.22; published 27.05.22.

Please cite as:

Asbjørnsen RA, Hjelvesæth J, Smedsrød ML, Wentzel J, Ollivier M, Clark MM, van Gemert-Pijnen JEW, Solberg Nes L. Combining Persuasive System Design Principles and Behavior Change Techniques in Digital Interventions Supporting Long-term Weight Loss Maintenance: Design and Development of eCHANGE

JMIR Hum Factors 2022;9(2):e37372

URL: <https://humanfactors.jmir.org/2022/2/e37372>

doi: [10.2196/37372](https://doi.org/10.2196/37372)

PMID: [35622394](https://pubmed.ncbi.nlm.nih.gov/35622394/)

©Rikke Aune Asbjørnsen, Jøran Hjelvesæth, Mirjam Lien Smedsrød, Jobke Wentzel, Marianne Ollivier, Matthew M Clark, Julia E W C van Gemert-Pijnen, Lise Solberg Nes. Originally published in *JMIR Human Factors* (<https://humanfactors.jmir.org>), 27.05.2022. This is an open-access article distributed under the terms of the Creative Commons Attribution License (<https://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in *JMIR Human Factors*, is properly cited. The complete bibliographic information, a link to the original publication on <https://humanfactors.jmir.org>, as well as this copyright and license information must be included.

Review

Chinese Americans' Use of Patient Portal Systems: Scoping Review

Katharine Lawrence^{1*}, MD, MPH; Stella Chong^{2*}, BA; Holly Krelle^{3*}, MPhil; Timothy Roberts^{4*}, MPH, MLS; Lorna Thorpe⁵, MPH, PhD; Chau Trinh-Shevrin², DrPH; Stella Yi², MPH, PhD; Simona Kwon², MPH, DrPH

¹Healthcare Innovation Bridging Research, Informatics, and Design (HiBRID) Lab, Department of Population Health, New York University Grossman School of Medicine, New York, NY, United States

²Section for Health Equity, Department of Population Health, NYU Grossman School of Medicine, New York, NY, United States

³Division of Healthcare Delivery Services, Department of Population Health, NYU Grossman School of Medicine, New York, NY, United States

⁴NYU Health Sciences Library, NYU Grossman School of Medicine, New York, NY, United States

⁵Division of Epidemiology, Department of Population Health, NYU Grossman School of Medicine, New York, NY, United States

*these authors contributed equally

Corresponding Author:

Katharine Lawrence, MD, MPH

Healthcare Innovation Bridging Research, Informatics, and Design (HiBRID) Lab

Department of Population Health

New York University Grossman School of Medicine

227 E 30th St

6th Floor

New York, NY, 10016

United States

Phone: 1 6465013488

Email: katharine.lawrence@nyulangone.org

Abstract

Background: Electronic patient portals are increasingly used in health care systems as communication and information-sharing tools and show promise in addressing health care access, quality, and outcomes. However, limited research exists on portal use patterns and practices among diverse patient populations, resulting in the lack of culturally and contextually tailored portal systems for these patients.

Objective: This study aimed to summarize existing evidence on the access and use patterns, barriers, and facilitators of patient portals among Chinese Americans, who represent a growing patient population in the United States with unique health care and health technology needs.

Methods: The authors conducted a literature search using the PRISMA Protocol for Scoping Reviews (Preferred Reporting Items for Systematic Reviews and Meta-Analyses-ScR) for extracting articles published in major databases (MEDLINE, Embase, and PsycINFO) on patient portals and Chinese Americans. Authors independently reviewed the papers during initial screening and full-text review. The studies were analyzed and coded for the study method type, sample population, and main outcomes of interest.

Results: In total, 17 articles were selected for inclusion in the review. The included articles were heterogeneous and varied in their study aims, methodologies, sample populations, and outcomes. Major findings identified from the articles include variable patterns of portal access and use among Chinese Americans compared to other racial or ethnic groups, with limited evidence on the specific barriers and facilitators for this group; a preference for cross-sectional quantitative tools such as patient surveys and electronic health record-based data over qualitative or other methodologies; and a pattern of aggregating Chinese American-related data into a larger Asian or Asian American designation.

Conclusions: There is limited research evaluating the use patterns, experiences, and needs of Chinese Americans who access and use patient portal systems. Existing research is heterogeneous, largely cross-sectional, and does not disaggregate Chinese Americans from larger Asian demographics. Future research should be devoted to the specific portal use patterns, preferences, and needs of Chinese Americans to help ensure contextually appropriate and acceptable design and implementation of these digital health tools.

KEYWORDS

patient portal; electronic health records; personal health records; ehealth; health equity; digital divide; Chinese Americans; Asian Americans

Introduction

The expansion of health information technology (HIT) has provided patients with tools to proactively access their health information, self-manage chronic conditions, and communicate directly with providers [1]. In particular, electronic patient portals—which are secure internet-based platforms or websites that provide patients with 24-hour access to their personal health information—have emerged as a common communication and information-sharing tool for health care systems [2]. Patient portals offer a variety of features and functions for patients, such as the ability to access and review medical information, view lab and imaging results, schedule medical appointments and other visits, and interact with their health care providers [2-4]. Increasingly, these systems are directly integrated into electronic health record (EHR)-based platforms (eg, Epic MyChart or eClinical Works) or customer relationship management systems, as well as into the growing ecosystem of telehealth services. The COVID-19 pandemic expanded the use of patient portals as a facilitator of virtual health care and telemedicine, remote patient-provider communication, and monitoring [5-7]. Patient portals have demonstrated effectiveness in improving patient communication, engagement, and satisfaction [8,9], with some evidence on improvements in health outcomes [7,10] and lowered health care costs [6]. However, despite these benefits, adoption of and engagement with patient portals have varied, and significant disparities in the use of portal systems have been identified [2,11-18]. These disparities are shaped by individual, community, and structural factors such as social demographics (eg, socioeconomic status), health status (eg, disability diagnosis, chronic illness status), human-computer interface design challenges (eg, usability), and structural barriers (eg, lack of access to broadband internet).

Chinese Americans are a population frequently under- or mis-represented in health care, health delivery, and health research [19,20]. At roughly 5 million people, Chinese Americans comprise the largest subgroup of a heterogeneous community of Asian Americans and Pacific Islanders (AAPI), who themselves represent almost 10% of the US population [21-23]. Chinese American patients have distinct experiences interacting with the health care system [23,24], including care moderated by health technologies [25-27]. Although health disparities in this community have been identified and are mediated by factors such as language proficiency and immigration status [22,24], the details of these experiences are often obscured by problems with data collection and interpretation of health data that ignores the considerable heterogeneity and complexity of the AAPI designation [28].

To improve the effectiveness, acceptability, and use of digital health technologies such as patient portals among diverse communities, a better understanding of the use patterns and practices of the specific communities and their subgroups is

needed. This scoping review summarizes the existing evidence on patient portal perceptions, adoption, and use among Chinese Americans, and it highlights gaps and areas for further research on patient portal and digital health technology use among Chinese Americans and other diverse patient populations.

Methods

The aim of conducting a scoping review is to identify and broadly describe knowledge and research pertaining to a topic of interest as well as to identify trends, patterns, and gaps in the literature. Scoping reviews are ideal for research areas where the study question is broad or exploratory, there is limited literature on the topic, or study methodologies are diverse [29].

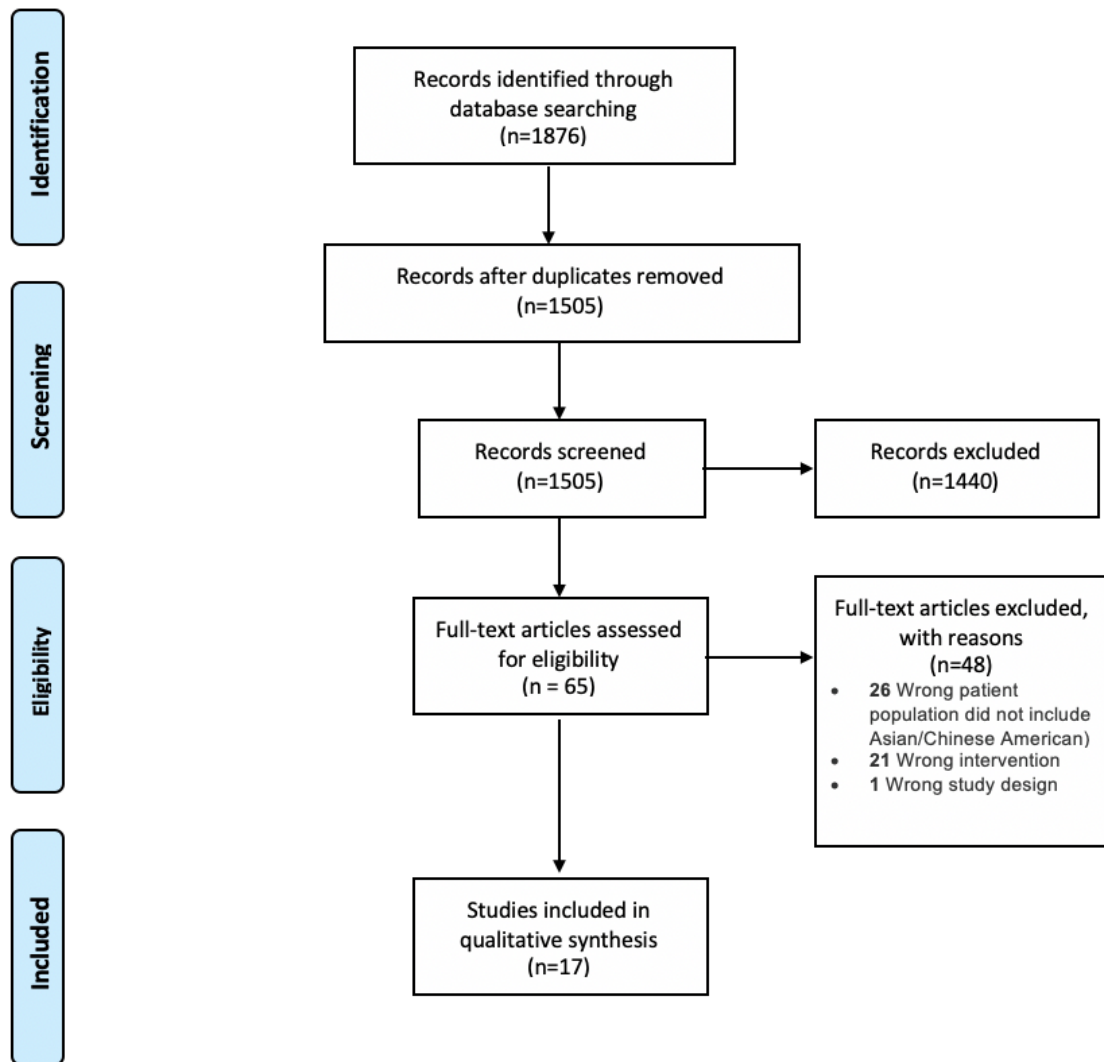
The review was conducted following the PRISMA Protocol for Scoping Reviews ((Preferred Reporting Items for Systematic Reviews and Meta-Analyses-ScR) [30]. In August 2020 and 2021, one of the coauthors (TR) who is an experienced medical librarian searched MEDLINE, Embase, and PsycINFO using the Ovid Platform and the Web of Science Core Collection. The search was not limited by language or publication date. Quantitative and qualitative studies that included primary data collection or data analysis were included; article types such as opinion pieces or letters to editors were excluded. The complete Ovid MEDLINE search strategy is available in [Multimedia Appendix 1](#).

US-based studies that described the inclusion and perceptions of Asian Americans (eg, Asians, Asian Americans, Chinese Americans, and Filipinos) toward electronic patient portals were included. Studies that identified Asian Americans only under the heading of “Other” without additional specificity were excluded. Patient portals were defined as web-based platforms that provided access to data from EHRs, including features such as medical histories, visit summaries, medication lists, as well as secure messaging features, access to educational resources, and appointment scheduling [19,20,31]. Studies that focused primarily on the delivery of “real-time interactive” remote clinical care using audio or video communication technology (eg, synchronous telemedicine) [32] were excluded, as these technologies often exist separately from patient portal communication systems or do not support key asynchronous features such as personal health data review by patients or remote monitoring. Studies exploring general health information literacy or information-seeking via digital resources (eg, the internet) in this group were also excluded.

After duplicates were removed, 1505 articles remained. Titles and abstracts were screened using Covidence software [33] by 2 independent reviewers (SKC and HK) for explicit or implicit mention or identification of Chinese Americans. Conflicts were resolved through discussion between the 2 reviewers until consensus was reached. When needed, consultation was sought from another coauthor (KL) to reach consensus. The full texts,

including tables, figures, and appendices, of 65 articles were reviewed following the same process. Ultimately, 17 articles were included in the analysis, as shown in Figure 1.

Figure 1. PRISMA Flowchart showing the screening and inclusion process of the studies. PRISMA: Preferred Reporting Items for Systematic Reviews and Meta-Analyses.



Results

Article Summaries

In total, 17 articles were selected for inclusion in the review. A summary of each article, including the study design, sample

information (including the level of Asian American population identified in the study), and key findings can be found in Tables 1 and 2.

Table 1. Characteristics of studies that include identifiable data specific to Chinese Americans.

Study	Objective	Research design/tools	Sample population/level of Chinese American granularity and location	Patient portal technology/feature	Relevant results
Ackerman et al (2017) [33]	To understand the implementation of patient portals in safety net health care systems striving to meet MU ^a criteria set by the Federal United States government	Mixed methods Rapid ethnography to assess MU, including interviews with providers and executives, informal focus groups with frontline staff, observations of patient portal sign-up procedures, and review of marketing materials and patient portal use. Administered modified version of the American Medical Association's Health IT ^b Readiness Survey Study tools: patient portal promotional flyers in English and Chinese at clinics; instructional video in Cantonese; language-congruent health staff available	5 California safety net health systems. Site 5 in Northern California, which serves 95% <i>non-native English-speaking Chinese immigrants</i> . Location: Northern California	Patient portal (NextGen) implementation strategies and efforts at Site 5 Patient portal features: medical history, test results, secure messaging, and appointment requests	Overall: Community health centers were motivated by MU incentives to increase patient portal enrollment and integrate portal-related work into clinic routines. Barriers to patient portal usage for patients: lack of internet access, lack of computer proficiency, discomfort with portal use, language barriers, fear of government surveillance, and preference for in-person interaction with providers. Specific to Site 5: Chinese American patients face language barriers in accessing the patient portal. "The (EHR ^c vendor) website isn't in their Chinese language... How were they going to get their patients to be able to utilize this?" Perception that clinic discouraged staff from promoting patient portal once MU threshold was reached.

Study	Objective	Research design/tools	Sample population/level of Chinese American granularity and location	Patient portal technology/feature	Relevant results
Gordon and Hornbrook (2016) [34]	To identify racial or ethnic and age disparities among older patients' use of patient portals and access to digital technology and devices for email and web-based health care management programs	Quantitative (cross-sectional, administrative data and survey) Study 1: Analyzed administrative data about patient portal account status and use from the KPNC ^d health plan Study 2: Mailed English survey questionnaire, from 2013 to 2014, to stratified random sample of Study 1's population	Study 1: English-speaking Chinese (n=6314), non-Hispanic White (n=183,565), Black (n=16,898), Latino (n=12,409), and Filipino (n=11,896) older patients aged 65 to 79 years. Study 2: same as Study 1 Location: Northern California	KPNC internet-based patient portal, kp.org, and other digital health technology and tools (eg, emails, text, computer, smartphones)	Compared to Black, Filipino, and Latino older patients, Chinese and non-Hispanic White older patients were more likely to be registered to use the patient portal and more likely to use portal functions. Chinese and non-Hispanic White older patients were more likely to access digital devices, internet, and email. They were also more likely to be willing to use digital technology to seek health information.
Gordon and Hornbrook (2018) [35]	To assess disparities by race/ethnicity and age on older patients' ability to engage with online health information and mobile health tools connected to their health system	Quantitative (cross-sectional, survey) Mailed English survey questionnaire, from November 2013 to February 2014 to members of the KPNC	Stratified random sample of 5420 <i>English-speaking</i> KPNC patients <i>Chinese</i> (n=500), non-Hispanic White (n=1420), African American/Black (n=1500), Hispanic/Latino (1500), and Filipino (n=500) Location: Northern California	Digital health technology and tools (eg, internet, computer, mobile phone, email, text, social media, apps)	Chinese and non-Hispanic White older patients have higher levels of access to digital tools, experience in performing a variety of web-based tasks, and belief in their ability to seek health information on the internet compared to Black, Latino, and Filipino peers. Chinese older people prefer having telephone appointments with health coaches and are less interested in reading about health topics on the internet. Chinese older people have the lowest level of interest in using health apps.

Study	Objective	Research design/tools	Sample population/level of Chinese American granularity and location	Patient portal technology/feature	Relevant results
Khoong et al (2020) [36]	To assess predictors of health technology use (eg, language preferences, smartphone ownership, type of clinic for health care)	Quantitative (cross-sectional, survey)	Nonrandom sample of 1027 participants <i>Chinese-speaking Chinese</i> (n=257); Spanish-speaking Latino (n=256); English-speaking non-Hispanic Black (n=514); English-speaking non-Hispanic White (n=43); and English-speaking Latino (115) Location: San Francisco, California	Digital health technology and tools for communication with clinicians (eg, email, text, phone apps, web-based health videos, and online health support groups)	Relative to English-speaking survey respondents, individuals who preferred the Chinese language had lower odds of texting or using an app to communicate with their clinician. There were no differences in using emails or watching web-based health videos. Language concordance was suggested as a major barrier.

^aMU: meaningful use.

^bIT: information technology.

^cEHR: electronic health record.

^dKPNC: Kaiser Permanente Northern California

Table 2. Characteristics of studies with aggregated Asian American data.

Study	Objective	Research design	Sample population/ level of Chinese American granularity and location	Focus	Relevant results
Ahlers-Schmidt and Nguyen (2013) [37]	To obtain parents' feedback and intention to use patient portals for their children's health records and concerns post the facilitated learning session	Quantitative (cross-sectional, survey)	Parents of patients. (N=65) White (n=26, 40%); Hispanic (n=14, 22%); <i>Asian</i> (n=9, 14%); African American (n=6, 9%); Mixed/other race (n=8, 12%) Location: Kansas	University of Kansas Pediatric Clinic's eClinical Works, an electronic medical record with a patient portal	Most parents did not know about the patient portal before the study demonstration. Parents expressed that patient portal was simple to use after demonstration. Parents liked portal functions such as viewing lab results and medical records; disliked need to make separate accounts for each child and the lack of a symptom checker function.
Dalrymple et al (2018) [38]	To assess parents' use of the internet for health information and parents' awareness of digital health technologies to obtain health information Screening questions assess parents' level of health literacy and interest in use of patient portals	Quantitative (cross-sectional, survey) Study tool: 26-question paper and pencil survey adapted from interview protocol designed from previous study	Total sample population of parents or adult caregivers of children and adolescents, N=270 <i>Asian</i> (1.9%); American Indian/Alaska Native (1.5%); Black/African American (38.1%); Hispanic/Latino (13.7%); Native Hawaiian/Pacific Islander (0.4%); White (40.7%); more than one race/ethnicity (4.4%); and Other (1.5%) Location: Unspecified large metropolitan area in eastern United States	Internet and patient portal	Most patients reported having access to the internet and using the internet to seek general and health information. Respondents expressed enthusiasm and interest in using a patient portal if it were available from their health care provider.
Foster and Kravowski (2019) [39]	To assess patient portal usage by ED ^a patients at an academic medical center using patient portal activation rates and rates of accessing diagnostic test results on patient portals	Quantitative (retrospective cohort, EHR ^b , and administrative data)	25,361 unique ED patients identified via EHR patient portal records <i>Asian</i> (n=451); African American/Black (n=2,254); White (n=20,637); Hispanic/Latino (n=1257); Other (n=762) Location: Iowa	UIHC ^c patient portal (MyChart), connected to EPIC EHR system	Highest rates of using the patient portal to view laboratory and radiology results were observed for younger female, proxies, <i>Asian</i> , and White patients. Activation rates were highest for <i>Asian</i> and White patients. Disparities were observed among teenagers, older adults, African American/Black, and Hispanic/Latino patients.

Study	Objective	Research design	Sample population/ level of Chinese American granularity and location	Focus	Relevant results
Goel et al (2011) [12]	To examine the enrollment in and use of patient portal at an academic medical center by race/ethnicity, gender, and age	Quantitative (cross-sectional, EHR and administrative data) Study tool: patients' use of EHR-based advice function and request for refills	Patients enrolled in the patient portal system, N=7088 <i>Asian</i> (n=142, 2%); White (n=3472, 49%); Black (n=1063, 15%); Latino (n=284, 4%); Other (n=851, 12%); Missing race/ethnicity (n=1347, 19%) Location: Chicago	Northwestern Medical Faculty Foundation's EHR patient portal	Significant disparities in patient portal enrollment by race/ethnicity were observed, but not by age or gender. White patients (74%) were more likely to enroll in patient portals compared to Black (55%), Latino (64%), and Asian (66%) patients. When adjusted for variables (eg, age, gender, income, education, and provider effects), the disparity between Asian and White patients was no longer statistically significant.
Graetz et al (2016) [40]	To assess socioeconomic disparities in patient portal use	Quantitative (cross-sectional, survey) Study tool: Administered paper-based survey mailed to participants; survey measures internet access, secure email use, care preference, sociodemographics, and health characteristics	Total study participants from KP-NC ^d , N=1041 White (n=617, 59.3%); <i>Asian</i> (n=145, 13.9%); Black (n=122, 11.7%), and Hispanic (n=128, 12.3%) Location: Northern California	Internet and email	Asian and Black respondents were more likely to rarely or never to use the internet (45.4% and 45.6%, respectively) compared to their White respondents. Asian participants (78%) preferred in-person care over telephone care compared to White patients (64%).
Ketterer et al (2013) [41]	To identify predictors of patient portal enrollment and activation among a pediatric primary care population	Quantitative (cross-sectional, EHR and administrative data) Study tool: primary care database, and enrollment in and use of a patient portal	Total sample population N=84,015 Black (n=35,286, 42%); <i>Asian</i> (n=2520, 3%); White (n=35,286, 42%); Hispanic (n=10,082, 12%); Other (n=9242, 11%); and Unknown (n=1680, 2%). Location: Delaware	Patient portal site, MyNemours	Adjusted odds of portal enrollment were lower for Asian respondents compared to White respondents. Once enrolled, there was no difference in portal activation between Asian respondents and White respondents. Study suggested language concordance as a major barrier.

Study	Objective	Research design	Sample population/ level of Chinese American granularity and location	Focus	Relevant results
Lyles et al (2013) [42]	To understand how patient-provider relationships influence patients' use of online patient portals and secure messaging	Quantitative (cross-sectional, survey)	Surveyed patients DISTANCE ^e Black (23%); Latino (16%); <i>East Asian (ie, Chinese, Japanese, Korean, or Vietnamese)</i> (10%); Filipino 12%); and Other (6%) Location: Northern California	KPNC's internet-based patient portal, kp.org	White and Latino individuals with higher trust in the providers were more likely to register on the patient portal. There was no relationship between trust in provider and patient portal use for Asian respondents.
Lyles et al (2016) [43]	To determine whether racial/ethnic minority patients' use of the patient portal's medication refill function has changed over time compared to White patients	Quantitative (EHR and administrative data) Study tool: diabetic patients' use of EHR-based medication refill function	White (58%); <i>Asian</i> (10%); Latino (9%); Filipino (9%); Black (7%); and Mixed/other (9%) Location: Northern California	KPNC's internet-based patient portal, kp.org	Asian were not less likely to exclusively use refill functions than other ethnic groups. Adherence to medication refills improved over time for all ethnic groups, but there was no significant difference between ethnicities. Usability and accessibility were identified as barriers to portal registration.
Miles et al (2016) [44]	To measure and evaluate the frequency at which patients use the patient portal to view online radiology reports	Quantitative (cross-sectional, EHR and administrative data) Study tool: patient interactions with portal features (eg, radiology, laboratory, and clinical notes) and sociodemographic factors	<i>Asian or Pacific Islander</i> (n=6376, 10.4%); American Indian or Alaska Native (n=522, 0.8%); Black or African American (n=3817, 6.2%); Hispanic or Latino (n=1850, 3%); White (n=44,163, 72.25); and Other/more than one race (n=675, 1.1%); and Unknown (n=3728, 6.1%) Location: Seattle, Washington	UW's ^f patient portal system, UW eCare web portal	Asian respondents were more likely than White patients to view their radiology reports. Older patients, primary non-English speakers, and those with non-commercial insurance viewed reports at lower rates. Concerns identified in the study include loss of patient confidentiality, health information inaccuracy, and disruption of patient-physician relationship.
Patel et al (2011) [45]	To determine low-income, ethnically diverse consumers' attitudes and beliefs toward HIE ^g and use of HIE via PHRs ^h and to identify factors that impact consumers' support for providers' use of HIE and their own personal use of PHRs	Quantitative (cross-sectional, survey) Study tool: survey adapted from previously validated national surveys. Survey was translated into Spanish, Russian, and Mandarin Chinese	BHIX ⁱ 's patients White (n=36, 74%); <i>Asian</i> (n=57, 28%); African American (n=20; 10%); and Other (n=56, 27%). <i>Spoke Chinese at home</i> (n=42, 20%) Location: New York City, New York	EHRs, internet, HIE, and PHRs	Compared to other racial/ethnic groups in the study, Asian Americans indicated lower levels of support for HIE (48%) and lower levels of potential PHR usage (67%).

Study	Objective	Research design	Sample population/ level of Chinese American granularity and location	Focus	Relevant results
Sarkar et al (2010) [46]	To examine whether use of an internet-based patient portal differed between English-speaking patients with limited health literacy and English-speaking patients with adequate health literacy	Quantitative (cross-sectional, survey) Study tool: DISTANCE study was conducted in English, Spanish, Cantonese, Mandarin, and Tagalog	Total of 14,201 surveyed participants from DISTANCE study Non-Hispanic White (n=3957, 28%); Latino (n=1923, 14%); African American (n=2899, 21%); Asian (n=1253, 9%); Filipino (n=1624, 12%); Other (n=2446, 17%) Location: Northern California	KPNC's internet-based patient portal, kp.org	Study did not find increased risk of not signing onto the patient portal for Asian Americans compared to African American, Latino, and Filipino respondents. Asian Americans had lower rates of never using patient portal functions including lab result viewing, medication refills, email, and scheduling appointments. Health literacy was identified as a barrier to portal activity.
Sarkar et al (2011) [47]	To examine portal use habits via the frequency at which participants requested a password for the patient portal, the proportion of participants who activated their accounts by changing the default password, and the proportion of participants who login to their accounts using their personal, customized password	Quantitative (cross-sectional, EHR and administrative data) DISTANCE study was conducted in English, Spanish, Cantonese, Mandarin, and Tagalog	Total of 14,201 surveyed participants from DISTANCE Study Non-Hispanic White (n=3957, 28%); Latino (n=1923, 14%); African American (n=2899, 21%); Asian (n=1253, 9%); Filipino (n=1624, 12%); Other (n=2446, 17%) Location: Northern California	KPNC's internet-based patient portal, kp.org	Asian American (53%) and White (51%) participants were more likely than their African American (31%), Latino (34%), and Filipino (32) counterparts to request a password for the internet-based patient portal and to login to the patient portal after requesting a password. Older adults with less educational attainment were less likely to register and use the patient portal.

Study	Objective	Research design	Sample population/ level of Chinese American granularity and location	Focus	Relevant results
Tieu et al (2017) [48]	To measure participants' satisfaction with use of patient portal	Mixed methods (cross-sectional, usability testing and survey) Study tool: Conducted English language performance testing and think-aloud interviews with participants and administered survey to participants	Total of 25 English-speaking (23 patients and 2 caregivers) participants. African American (n=9, 36%); White (n=6, 24%); Hispanic (n=2, 8%); <i>Asian or Pacific Islander</i> (n=5, 20%); and Other (n=3, 12%) Location: San Francisco, California	RFPC's ^j patient portal, MYSFHEALTH	Participants with limited health literacy, including Asian and Pacific Islander patients were more likely to need assistance navigating the patient portal. Barriers to patient portal use for participants with limited health literacy include (1) lack of basic computer skills; (2) routine computer use challenges despite basic knowledge of computers; (3) difficulty reading, writing, and understanding language; and (4) difficulty understanding and applying medical information from the internet and patient portal.

^aED: emergency department.

^bEHR: electronic health record.

^cUIHC: University of Iowa Hospitals and Clinics.

^dKPNC: Kaiser Permanente Northern California.

^eDISTANCE: Diabetes Study of North California.

^fUW: University of Washington.

^gHIE: health information exchange.

^hPHRs: personal health records.

ⁱBHIX: Brooklyn Health Information Exchange.

^jRFPC: Richard H. Fine People's Clinic.

The included articles varied in terms of the study methodology, sample population, data collection methodology, and geographic area within the United States. Among these, 10 were from populations in California [34-36,40,42,43,45-48], with 5 from the Kaiser Permanente health system [35,36,42,43,46]. Further, 3 studies used a shared database—the Diabetes Study of Northern California (DISTANCE)—to analyze portal-related outcomes [40,47,48]. Of the data collection tools described in these studies, 5 studies indicated they were available and conducted in Chinese (eg, Mandarin or Cantonese) [12,34,43,47,48].

Overall, the articles described heterogeneous results among varied patient populations, health conditions, and care settings. Few clear themes emerged and results specific to Asian American subgroups such as Chinese Americans were not identified. In general, the authors were able to identify the following major themes and trends from the results.

1. Chinese Americans demonstrate variable patterns of patient portal access and use as compared to other demographics,

particularly racial or ethnic groups; exploration of the specific contexts of use, including barriers and facilitators, is limited.

2. Most studies employed cross-sectional, quantitative tools to assess patient portal use patterns and practices, including patient surveys and EHR-based data that measure portal activity (eg, logins and click-throughs); neither longitudinal nor significant qualitative research studies were conducted to validate or further explore nuances in findings specific to Chinese Americans.
3. Despite the heterogeneity of the populations included in AAPI designation, studies exploring patient portals do not disaggregate Asian and Asian American study populations into Chinese Americans and other subgroups.

Findings Specific to Chinese Americans

Only 4 studies [33-36] specifically disaggregated Chinese American populations (Table 1). All 4 of these were from California. Among these, 3 [34-36] were primarily based around surveys, and 1 [33] was based on rapid ethnography, mostly focusing on understanding the barriers to accessing patient

portals. Barriers reported included language barriers, lack of internet access or computer proficiency, fear of government surveillance, and a preference for in-person interaction. Further, 2 of the studies [34,35] found that Chinese patients were more likely than other non-White groups to register and use internet-based portals, and 1 [40] found that relative to English-speaking respondents, people who preferred the Chinese language were less likely to send text messages or use an app to contact their clinician.

Chinese Americans Demonstrate Variable Patterns of Patient Portal Access and Use Compared to Other Racial or Ethnic Groups

This represents a finding in the data across studies, with some demonstrating lower rates of use and others demonstrating higher rates and rates comparable to White patients. In a study on the use of the Northwestern Medical Faculty Foundation's electronic patient portal [44], the authors found that once variables such as age, gender, education, income, and provider effects were adjusted, there was no disparity between the enrollments of Asian American and White patients on the patient portal. In another study of Chinese American older adults in Kaiser Permanente, Northern California [35], the authors found that non-Hispanic White and Chinese American older adults were more likely than other racial or ethnic groups to register for using the portal and its functions such as sending messages, viewing lab results, or ordering prescription refills. Other studies showed lower use and lower motivation to use digital health technology among Chinese Americans. In their study examining patients' patterns of texting and communication with their clinicians via apps, Khoong et al [36] found that individuals who preferred to use Chinese language had lower odds of texting or using an app to communicate with their clinicians compared to English-speaking survey respondents. In a study assessing older patients' readiness to use eHealth tools, researchers found that Chinese American patients had the lowest level of interest in using patient portal technology among all the racial or ethnic groups in the study, though their experience of using the internet was similar to that of non-Hispanic White patients [36]. In their assessment of attitudes toward health information exchanges (HIEs) and personal health records (PHRs), Patel et al [45] found that Asian Americans were less likely than other racial or ethnic groups to support the use of PHR technology.

Identified studies provided limited evidence on the barriers faced by Chinese Americans in using patient portals. For individuals, the main reported barrier was language congruency with the portal or related technologies, or English language proficiency. In a mixed methods study evaluating the implementation of meaningful use at community health centers in California, Ackerman et al [33] noted that many patients could not read English and that even if communication with care providers could be conducted in Chinese, most EHR features (including records, test results, and communication tools like the patient portal) were exclusively in English. The authors also noted concerns among some Chinese Americans about government surveillance, particularly among patients who were undocumented or had concerns regarding their immigration status. Additional individual-level barriers identified in the

studies included issues of usability and accessibility of the portal tool [43], concerns around confidentiality and privacy [38], low health literacy, [48], and digital literacy [45]. Conversely, in a study assessing the influence of patient-provider relationships on patient portal and messaging usage, Lyles et al [42] found that although trust in providers was correlated with registration for portals by White and Latinx patients, this was not the case for Asian patients.

Identified community and structural barriers were largely related to clinic-level resources and included the clinical staff's ability to support patients' engagement in patient portal technology and the paucity of language-congruent support services. In their rapid ethnography with clinical staff in safety net hospital-affiliated practices, Ackerman et al [33] reported challenges related to providers and staff members having limited time and skills to coach patients in using the patient portal, and concerns regarding meaningful use metrics that prioritize outcomes such as portal sign-up rather than sustained use. The researchers also identified disruptions to clinical workflows and increased administrative burden as barriers to effective implementation and use of EHR-related tools. In 3 studies, access to digital technology and infrastructure such as the internet was associated with higher rates of patient portal access and use by Chinese and Asian American patients [35-37].

Most Studies Employed Cross-sectional Quantitative Tools to Assess Patient Portal Use Patterns and Practices, Including Patient Surveys and EHR-Based Data That Measure Portal Activity

Among the 17 studies, 8 employed survey-based, numeric (eg, Likert scale) data collection tools disseminated using either digital tools (eg, email) or in person. Survey question areas ranged from portal familiarity and general perspectives to personal experiences, feature preferences, and self-reporting of details on use habits [12,35-37,39,40,43,47]. The remaining studies used either administrative information-based EHRs or associated databases. Furthermore, 6 studies conducted primary EHR-based analyses to identify patterns and trends in portal-based activities [38,41,44,46,48,49]. Key EHR- and portal-based measures reported by researchers included patient portal registrations [35,44], logins and appointment booking [47-49], medication refill requests [46], viewing of results and reports (eg, radiology reports) [38,41,47], and texting and other forms of communication with clinicians [40,42]. These activities were analyzed for frequency and other patterns, and they were often compared among demographics such as age, race or ethnicity, sex or gender, income level, insurance status, and language. Key themes in the survey questions included actual and expected use of different features, concerns and barriers related to using portals, and confidence in the ability to use portals and understand health information shared through these portals. Most of these measures are applied cross-sectionally, and there is neither longitudinal nor significant qualitative research to validate or further explore nuances in findings specific to Chinese Americans or other Asian American subgroups. No studies included measures of associated health outcomes.

Despite the Heterogeneity of the Populations Included in AAPI Designation, Studies Exploring Patient Portals Largely do not Disaggregate Asian and Asian American Study Populations

Of the 17 studies included in this review, only 4 specifically disaggregate or discuss Chinese Americans [34-36,40]. The remaining studies generally refer to “Asian Americans” or “Asians,” with only indirect references to over 20 unique ethnic subgroups included in that designation or otherwise included in the study sample, data collection, or analysis. For example, Chinese-speaking patients were occasionally mentioned in the text or tables of these studies [12,34,43,47,48] but not included in any multivariate analyses as a separate category. In these studies, it was inferred that Chinese American patients were included via references to the languages of the data collection instruments (eg, Mandarin or Cantonese) or the study database being used for analysis. No studies specifically or exclusively evaluated Chinese Americans’ attitudes toward, perceptions about, or use of patient portal technology.

Discussion

Principal Findings

This scoping review highlights the extremely limited research on the use patterns, experiences, and needs of Chinese Americans who access and use patient portal systems for their health care. The identified studies were heterogenous in their approaches and outcomes, making generalizable trends in the data difficult to identify, although we were able to identify some patterns in the research methodologies and data collection tools across studies. By and large, the existing studies have focused on the identification of varying portal use patterns among racial, ethnic, and other demographics, and their correlative predictors such as age, primary language, or health literacy. Overall, the studies obtained mixed findings regarding the rates of portal usage by Chinese Americans when compared to other populations, with some indicating lower rates of portal adoption and use when compared to White patients and others finding comparable rates. We were unable to identify trends more granularly in terms of portal access within Chinese American subgroups (eg, women, geographic populations) due to limitations in the available data. We identified individual- and system-level factors that contributed to use patterns, as well as barriers to access and usage. Relevant individual-level factors included English language proficiency and language congruency with portal technology; health literacy; perceived usability and usefulness of the technology; and trust in provider relationships, privacy, and confidentiality. Relevant system-level factors included clinical resource and capacity limitations, and access to digital tools such as email and the internet. Studies tended to be cross-sectional and quantitative in nature, with minimal exploration of longitudinal trends in use patterns or practices, qualitative aspects, or correlation with health outcomes. Finally, we identified a pattern of data aggregation practices that tended to combine and compare Asian Americans as a larger demographic group to other racial or ethnic groups, rather than identifying data at the level of Chinese Americans or other subgroups. This practice had the effect of generalizing learning

across Asian Americans, thus providing limited insight into the experiences of Asian subgroups of different ethnicities, languages, and religious affiliations, among other factors.

To our knowledge, this is the first study to evaluate the patient portal use patterns and needs of Chinese Americans. Prior research has explored various features of patient portal activity, use, and experience in other clinical contexts, including among Black and Latinx communities and vulnerable populations such as the older people and those with disabilities [50-52]. A comprehensive review of interventions to increase patient portal use in “vulnerable populations” by Grossman et al in 2019 [4] identified 18 studies evaluating the impact of interventions designed to increase portal use or reduce disparities in use. The authors noted that most studies focused on individual-level interventions such as patient education and training and identified a lack of interventions or programs targeting tool- (eg, patient portal interfaces or features), community-, organizational-, or system-level factors to improve portal adoption and use [4]. This is also supported by the findings of the study led by Antonio et al [52] that explored patient portal research through the lens of health equity and identified a varying and often superficial level of interest in portal technology among underserved groups by researchers and an underemphasis on the systemic factors influencing patient portal access and use among diverse communities. Although comprehensive, these reviews included limited information on the needs, use patterns, or potential interventions for specific vulnerable groups, particularly among racial or ethnic demographics; as observed in our findings, data on race and ethnicity included in these reviews often excluded Asian Americans or did not identify Asian American subgroups. Though our study includes some of the articles referenced by these reviews, our focus on Chinese and Asian American subgroups provides additional specificity to the overall literature on patient portals and exposes existing challenges in identifying and applying appropriately tailored solutions to technical problems for undifferentiated “vulnerable” patients.

The findings of this study have important implications for the design and deployment of patient portals and other digital health tools (eg, EHRs, mobile health apps) as well as for the study of health technology usage among Chinese Americans, Asian subgroups, and other diverse or vulnerable patient populations. Overall, there is need for a more granular study focusing on the use of digital health technology by diverse communities to elucidate key differences in their needs, preferences, and constraints. Participatory design frameworks that incorporate diverse stakeholders to identify and address specific needs, preferences, and concerns regarding health care technologies can help inform more effective and sustainable implementation of these tools in clinical practice. Frameworks and methodologies that explicitly address digital health disparities and digital health equity, such as the equity-centered design framework [53] and the digital health equity framework [54], can additionally help identify and overcome structural barriers such as access to digital infrastructure or institutional racism. At the same time, there is a need for clearer definitions and more granular breakdowns of populations included in data collection and data publication processes to better inform

appropriate, targeted recommendations for diverse communities. Critically, the use of aggregate data as a proxy for subsets of Asian American patients obscures differences in patient- and community-level experiences or needs and conflates the experiences of minority communities within that population. This problematic practice has been well documented, and efforts are in place to address it in research and clinical practice [27,55-57]. Health informaticists and technology researchers can be change leaders in this area by applying well-established design practices such as user stories, personas, and customer segmentation to clearly identify the needs of patient users, including those that are defined by a specific cultural identity or intersections of identities [58,59].

There are several limitations to this study. We included only major databases (PubMed and Embase) and did not include unpublished or gray literature. We also limited our inclusion criteria to articles published only in English, excluding Chinese language biomedical databases such as the China National Knowledge Infrastructure. We further included only those articles focusing on populations in the United States. These criteria were established to ensure a focused review of our target community of interest, namely Chinese Americans, engaging with relatively similar health care delivery models and HIT technology. However, this may have resulted in the exclusion of relevant articles, particularly those published in Chinese language journals. Additionally, although the term “patient portal” included in our search string is broadly used, our search may have missed studies that incorporated portals, portal-like systems (eg, PHRs), or portal features without explicitly identifying them. We attempted to address this by performing

a series of web-based searches (Google) and manual searches to identify articles using variable terms that could meet our inclusion criteria. Finally, our study did not systematically evaluate the quality of the data presented in the included studies beyond an assessment of the study design and the level of racial or ethnic granularity among Asian Americans; moreover, we did not evaluate the bias in these studies. Future areas of research may include expanded language contexts and further quality and bias evaluations.

Conclusions

There is limited research dedicated to understanding the use patterns, experiences, and needs of Chinese Americans who access and use patient portal systems for their health care. Most of the research in this area focuses on disparities in use and access across the aggregated racial and ethnic demographic of Asian Americans, potentially obscuring important differences among and between the diverse and heterogeneous populations that comprise this designation. Studies are also overwhelmingly quantitative, focused on surveys and administrative data from portal systems, and they lack longitudinal data. Future research should focus specifically on Chinese Americans and prioritize performing more detailed longitudinal and qualitative evaluations to understand why specific communities of patients access and use portals in the ways that they do. A broader understanding of the diversity of health technology users in general can help ensure that these tools are applicable and acceptable to all patients, including the most vulnerable, and do not contribute to disparities in health access, equity, or outcomes.

Acknowledgments

This publication is supported by the National Institute on Minority Health and Health Disparities-National Institutes of Health (grant U54MD000538). The views expressed are those of the authors and do not necessarily represent the official position of the funding organization.

Conflicts of Interest

None declared.

Multimedia Appendix 1

Ovid MEDLINE search strategy.

[[PNG File , 282 KB](#) - [humanfactors_v9i2e27924_app1.png](#)]

References

1. Kruse CS, Beane A. Health information technology continues to show positive effect on medical outcomes: systematic review. *J Med Internet Res* 2018 Feb;20(2):e41 [[FREE Full text](#)] [doi: [10.2196/jmir.8793](#)] [Medline: [29402759](#)]
2. Ancker JS, Barrón Y, Rockoff ML, Hauser D, Pichardo M, Szerencsy A, et al. Use of an electronic patient portal among disadvantaged populations. *J Gen Intern Med* 2011 Oct;26(10):1117-1123 [[FREE Full text](#)] [doi: [10.1007/s11606-011-1749-y](#)] [Medline: [21647748](#)]
3. Otokiti A, Williams KS, Warsame L. Impact of digital divide on the adoption of online patient portals for self-motivated patients. *Healthc Inform Res* 2020 Jul;26(3):220-228 [[FREE Full text](#)] [doi: [10.4258/hir.2020.26.3.220](#)] [Medline: [32819040](#)]
4. Grossman L, Masterson Creber RM, Benda N, Wright D, Vawdrey D, Ancker J. Interventions to increase patient portal use in vulnerable populations: a systematic review. *J Am Med Inform Assoc* 2019 Aug;26(8-9):855-870 [[FREE Full text](#)] [doi: [10.1093/jamia/ocz023](#)] [Medline: [30958532](#)]

5. Grange ES, Neil EJ, Stoffel M, Singh AP, Tseng E, Resco-Summers K, et al. Responding to COVID-19: the UW medicine information technology services experience. *Appl Clin Inform* 2020 Mar;11(2):265-275 [FREE Full text] [doi: [10.1055/s-0040-1709715](https://doi.org/10.1055/s-0040-1709715)] [Medline: [32268390](https://pubmed.ncbi.nlm.nih.gov/32268390/)]
6. Darcourt JG, Aparicio K, Dorsey PM, Ensor JE, Zsigmond EM, Wong ST, et al. Analysis of the implementation of telehealth visits for care of patients with cancer in Houston during the COVID-19 pandemic. *JCO Oncol Pract* 2021 Jan;17(1):e36-e43. [doi: [10.1200/OP.20.00572](https://doi.org/10.1200/OP.20.00572)] [Medline: [33026951](https://pubmed.ncbi.nlm.nih.gov/33026951/)]
7. Sherwin J, Lawrence K, Gragnano V, Testa PA. Scaling virtual health at the epicentre of coronavirus disease 2019: a case study from NYU Langone Health. *J Telemed Telecare* 2020 Jul;1-6. [doi: [10.1177/1357633X20941395](https://doi.org/10.1177/1357633X20941395)] [Medline: [32686555](https://pubmed.ncbi.nlm.nih.gov/32686555/)]
8. Ammenwerth E, Schnell-Inderst P, Hoerbst A. The impact of electronic patient portals on patient care: a systematic review of controlled trials. *J Med Internet Res* 2012 Nov;14(6):e162 [FREE Full text] [doi: [10.2196/jmir.2238](https://doi.org/10.2196/jmir.2238)] [Medline: [23183044](https://pubmed.ncbi.nlm.nih.gov/23183044/)]
9. Irizarry T, DeVito DA, Curran CR. Patient portals and patient engagement: a state of the science review. *J Med Internet Res* 2015 Jun;17(6):e148 [FREE Full text] [doi: [10.2196/jmir.4255](https://doi.org/10.2196/jmir.4255)] [Medline: [26104044](https://pubmed.ncbi.nlm.nih.gov/26104044/)]
10. Goldzweig CL, Orshansky G, Paige NM, Towfigh AA, Haggstrom DA, Miake-Lye I, et al. Electronic patient portals: evidence on health outcomes, satisfaction, efficiency, and attitudes: a systematic review. *Ann Intern Med* 2013 Nov;159(10):677-687. [doi: [10.7326/0003-4819-159-10-201311190-00006](https://doi.org/10.7326/0003-4819-159-10-201311190-00006)] [Medline: [24247673](https://pubmed.ncbi.nlm.nih.gov/24247673/)]
11. Lober W, Zierler B, Herbaugh A, Shinstrom SE, Stolyar A, Kim EH, et al. Barriers to the use of a personal health record by an elderly population. *AMIA Annu Symp Proc* 2006:514-518 [FREE Full text] [Medline: [17238394](https://pubmed.ncbi.nlm.nih.gov/17238394/)]
12. Goel MS, Brown TL, Williams A, Cooper AJ, Hasnain-Wynia R, Baker DW. Patient reported barriers to enrolling in a patient portal. *J Am Med Inform Assoc* 2011 Dec;18(Supplement 1):i8-i12 [FREE Full text] [doi: [10.1136/amiajnl-2011-000473](https://doi.org/10.1136/amiajnl-2011-000473)] [Medline: [22071530](https://pubmed.ncbi.nlm.nih.gov/22071530/)]
13. Roblin DW, Houston TK, Allison JJ, Joski PJ, Becker ER. Disparities in use of a personal health record in a managed care organization. *J Am Med Inform Assoc* 2009 Sep;16(5):683-689 [FREE Full text] [doi: [10.1197/jamia.M3169](https://doi.org/10.1197/jamia.M3169)] [Medline: [19567790](https://pubmed.ncbi.nlm.nih.gov/19567790/)]
14. Yamin CK, Emani S, Williams DH, Lipsitz SR, Karson AS, Wald JS, et al. The digital divide in adoption and use of a personal health record. *Arch Intern Med* 2011 Mar;171(6):568-574. [doi: [10.1001/archinternmed.2011.34](https://doi.org/10.1001/archinternmed.2011.34)] [Medline: [21444847](https://pubmed.ncbi.nlm.nih.gov/21444847/)]
15. Ancker JS, Hafeez B, Kaushal R. Socioeconomic disparities in adoption of personal health records over time. *Am J Manag Care* 2016 Aug;22(8):539-540 [FREE Full text] [Medline: [27541700](https://pubmed.ncbi.nlm.nih.gov/27541700/)]
16. Kim E, Stolyar A, Lober WB, Herbaugh AL, Shinstrom SE, Zierler BK, et al. Challenges to using an electronic personal health record by a low-income elderly population. *J Med Internet Res* 2009 Oct;11(4):e44 [FREE Full text] [doi: [10.2196/jmir.1256](https://doi.org/10.2196/jmir.1256)] [Medline: [19861298](https://pubmed.ncbi.nlm.nih.gov/19861298/)]
17. Atreja A, Mehta N, Miller D, Moore S, Nichols K, Miller H, et al. One size does not fit all: using qualitative methods to inform the development of an Internet portal for multiple sclerosis patients. *AMIA Annu Symp Proc* 2005:16-20 [FREE Full text] [Medline: [16778993](https://pubmed.ncbi.nlm.nih.gov/16778993/)]
18. Kruse RL, Koopman RJ, Wakefield BJ, Wakefield DS, Keplinger LE, Canfield SM, et al. Internet use by primary care patients: Where is the digital divide? *Fam Med* 2012 May;44(5):342-347 [FREE Full text] [Medline: [23027117](https://pubmed.ncbi.nlm.nih.gov/23027117/)]
19. Shah NS, Kandula NR. Addressing Asian American misrepresentation and underrepresentation in research. *Ethn Dis* 2020 Jul;30(3):513-516. [doi: [10.18865/ed.30.3.513](https://doi.org/10.18865/ed.30.3.513)] [Medline: [32742157](https://pubmed.ncbi.nlm.nih.gov/32742157/)]
20. Đoàn LN, Takata Y, Sakuma KK, Irvin VL. Trends in clinical research including Asian American, Native Hawaiian, and Pacific Islander participants funded by the US National Institutes of Health, 1992 to 2018. *JAMA Netw Open* 2019 Jul;2(7):e197432. [doi: [10.1001/jamanetworkopen.2019.7432](https://doi.org/10.1001/jamanetworkopen.2019.7432)] [Medline: [31339543](https://pubmed.ncbi.nlm.nih.gov/31339543/)]
21. Infographic—percentage of Asian Americans with limited English proficiency (2017). AAPI Data. 2020 May 28. URL: <https://aapidata.com/infographic-limited-english-2-2/> [accessed 2021-07-01]
22. Asian American and Pacific Islander heritage month: May 2020. The United States Census Bureau. URL: <https://www.census.gov/newsroom/facts-for-features/2020/aian.html> [accessed 2021-07-01]
23. Ma GX. Barriers to the use of health services by Chinese Americans. *J Allied Health* 2000 Jan;29(2):64-70. [Medline: [10874332](https://pubmed.ncbi.nlm.nih.gov/10874332/)]
24. Ngo-Metzger Q, Hayes GR, Yunan C, Cygan R, Garfield CF. Improving communication between patients and providers using health information technology and other quality improvement strategies: focus on Asian Americans. *Med Care Res Rev* 2010 Oct;67(5 Suppl):231S-245S. [doi: [10.1177/1077558710375432](https://doi.org/10.1177/1077558710375432)] [Medline: [20675345](https://pubmed.ncbi.nlm.nih.gov/20675345/)]
25. Li W, Zhong J. A review of technology-based interventions in improving type 2 diabetes management in Chinese Americans. *Asian Pac Isl Nurs J* 2019;4(1):20-33 [FREE Full text] [doi: [10.31372/20190401.1018](https://doi.org/10.31372/20190401.1018)] [Medline: [31037270](https://pubmed.ncbi.nlm.nih.gov/31037270/)]
26. Yeung A, Hails K, Chang T, Trinh N, Fava M. A study of the effectiveness of telepsychiatry-based culturally sensitive collaborative treatment of depressed Chinese Americans. *BMC Psychiatry* 2011 Sep;11(1):154. [doi: [10.1186/1471-244x-11-154](https://doi.org/10.1186/1471-244x-11-154)] [Medline: [21943315](https://pubmed.ncbi.nlm.nih.gov/21943315/)]
27. Holland AT, Palaniappan LP. Problems with the collection and interpretation of Asian-American health data: omission, aggregation, and extrapolation. *Ann Epidemiol* 2012 Jun;22(6):397-405 [FREE Full text] [doi: [10.1016/j.annepidem.2012.04.001](https://doi.org/10.1016/j.annepidem.2012.04.001)] [Medline: [22625997](https://pubmed.ncbi.nlm.nih.gov/22625997/)]

28. Levac D, Colquhoun H, O'Brien KK. Scoping studies: advancing the methodology. *Implement Sci* 2010 Sep;5:69 [FREE Full text] [doi: [10.1186/1748-5908-5-69](https://doi.org/10.1186/1748-5908-5-69)] [Medline: [20854677](https://pubmed.ncbi.nlm.nih.gov/20854677/)]
29. Tricco AC, Lillie E, Zarin W, O'Brien KK, Colquhoun H, Levac D, et al. PRISMA Extension for Scoping Reviews (PRISMA-ScR): checklist and explanation. *Ann Intern Med* 2018 Oct;169(7):467-473 [FREE Full text] [doi: [10.7326/M18-0850](https://doi.org/10.7326/M18-0850)] [Medline: [30178033](https://pubmed.ncbi.nlm.nih.gov/30178033/)]
30. Using telehealth to expand access to essential health services during the COVID-19 pandemic. Centers for Disease Control and Prevention. URL: <https://www.cdc.gov/coronavirus/2019-ncov/hcp/telehealth.html> [accessed 2021-04-15]
31. Powell KR. Patient-perceived facilitators of and barriers to electronic portal use: a systematic review. *Comput Inform Nurs* 2017 Dec;35(11):565-573. [doi: [10.1097/CIN.0000000000000377](https://doi.org/10.1097/CIN.0000000000000377)] [Medline: [28723832](https://pubmed.ncbi.nlm.nih.gov/28723832/)]
32. Better systematic review management. Covidence. URL: <https://www.covidence.org/> [accessed 2021-04-10]
33. Ackerman SL, Sarkar U, Tieu L, Handley MA, Schillinger D, Hahn K, et al. Meaningful use in the safety net: a rapid ethnography of patient portal implementation at five community health centers in California. *J Am Med Inform Assoc* 2017 Oct;24(5):903-912 [FREE Full text] [doi: [10.1093/jamia/ocx015](https://doi.org/10.1093/jamia/ocx015)] [Medline: [28340229](https://pubmed.ncbi.nlm.nih.gov/28340229/)]
34. Gordon NP, Hornbrook MC. Differences in access to and preferences for using patient portals and other eHealth technologies based on race, ethnicity, and age: a database and survey study of seniors in a large health plan. *J Med Internet Res* 2016 Mar;18(3):e50 [FREE Full text] [doi: [10.2196/jmir.5105](https://doi.org/10.2196/jmir.5105)] [Medline: [26944212](https://pubmed.ncbi.nlm.nih.gov/26944212/)]
35. Gordon NP, Hornbrook MC. Older adults' readiness to engage with eHealth patient education and self-care resources: a cross-sectional survey. *BMC Health Serv Res* 2018 Dec;18(1):220 [FREE Full text] [doi: [10.1186/s12913-018-2986-0](https://doi.org/10.1186/s12913-018-2986-0)] [Medline: [29587721](https://pubmed.ncbi.nlm.nih.gov/29587721/)]
36. Khoong EC, Rivadeneira NA, Hiatt RA, Sarkar U. The use of technology for communicating with clinicians or seeking health information in a multilingual urban cohort: cross-sectional survey. *J Med Internet Res* 2020 Apr;22(4):e16951 [FREE Full text] [doi: [10.2196/16951](https://doi.org/10.2196/16951)] [Medline: [32250280](https://pubmed.ncbi.nlm.nih.gov/32250280/)]
37. Ahlers-Schmidt CR, Nguyen M. Parent intention to use a patient portal as related to their children following a facilitated demonstration. *Telemed J E Health* 2013 Dec;19(12):979-981. [doi: [10.1089/tmj.2013.0041](https://doi.org/10.1089/tmj.2013.0041)] [Medline: [24050608](https://pubmed.ncbi.nlm.nih.gov/24050608/)]
38. Dalrymple PW, Rogers M, Zach L, Luberti A. Understanding Internet access and use to facilitate patient portal adoption. *Health Informatics J* 2018 Dec;24(4):368-378 [FREE Full text] [doi: [10.1177/1460458216675497](https://doi.org/10.1177/1460458216675497)] [Medline: [27821606](https://pubmed.ncbi.nlm.nih.gov/27821606/)]
39. Foster B, Krasowski MD. The use of an electronic health record patient portal to access diagnostic test results by emergency patients at an academic medical center: retrospective study. *J Med Internet Res* 2019 Jun;21(6):e13791. [doi: [10.2196/13791](https://doi.org/10.2196/13791)]
40. Graetz I, Gordon N, Fung V, Hamity C, Reed ME. The digital divide and patient portals: internet access explained differences in patient portal use for secure messaging by age, race, and income. *Med Care* 2016 Aug;54(8):772-779. [doi: [10.1097/MLR.0000000000000560](https://doi.org/10.1097/MLR.0000000000000560)] [Medline: [27314262](https://pubmed.ncbi.nlm.nih.gov/27314262/)]
41. Ketterer T, West DW, Sanders VP, Hossain J, Kondo MC, Sharif I. Correlates of patient portal enrollment and activation in primary care pediatrics. *Acad Pediatr* 2013 May;13(3):264-271. [doi: [10.1016/j.acap.2013.02.002](https://doi.org/10.1016/j.acap.2013.02.002)] [Medline: [23680344](https://pubmed.ncbi.nlm.nih.gov/23680344/)]
42. Lyles CR, Sarkar U, Ralston JD, Adler N, Schillinger D, Moffet HH, et al. Patient-provider communication and trust in relation to use of an online patient portal among diabetes patients: the diabetes and aging study. *J Am Med Inform Assoc* 2013 May;20(6):1128-1131 [FREE Full text] [doi: [10.1136/amiajnl-2012-001567](https://doi.org/10.1136/amiajnl-2012-001567)] [Medline: [23676243](https://pubmed.ncbi.nlm.nih.gov/23676243/)]
43. Lyles CR, Sarkar U, Schillinger D, Ralston JD, Allen JY, Nguyen R, et al. Refilling medications through an online patient portal: consistent improvements in adherence across racial/ethnic groups. *J Am Med Inform Assoc* 2015 Sep;23(e1):e28-e33 [FREE Full text] [doi: [10.1093/jamia/ocv126](https://doi.org/10.1093/jamia/ocv126)] [Medline: [26335983](https://pubmed.ncbi.nlm.nih.gov/26335983/)]
44. Miles RC, Hippe DS, Elmore JG, Wang CL, Payne TH, Lee CI. Patient access to online radiology reports: frequency and sociodemographic characteristics associated with use. *Acad Radiol* 2016 Sep;23(9):1162-1169. [doi: [10.1016/j.acra.2016.05.005](https://doi.org/10.1016/j.acra.2016.05.005)] [Medline: [27287715](https://pubmed.ncbi.nlm.nih.gov/27287715/)]
45. Patel VN, Dhopeswarkar RV, Edwards A, Barron Y, Likourezos A, Burd L, et al. Low-income, ethnically diverse consumers' perspective on health information exchange and personal health records. *Inform Health Soc Care* 2011 Dec;36(4):233-252. [doi: [10.3109/17538157.2011.554930](https://doi.org/10.3109/17538157.2011.554930)] [Medline: [21851182](https://pubmed.ncbi.nlm.nih.gov/21851182/)]
46. Sarkar U, Karter AJ, Liu JY, Adler NE, Nguyen R, Lopez A, et al. The literacy divide: health literacy and the use of an internet-based patient portal in an integrated health system-results from the diabetes study of northern California (DISTANCE). *J Health Commun* 2010 Aug;15(sup 2):183-196 [FREE Full text] [doi: [10.1080/10810730.2010.499988](https://doi.org/10.1080/10810730.2010.499988)] [Medline: [20845203](https://pubmed.ncbi.nlm.nih.gov/20845203/)]
47. Sarkar U, Karter AJ, Liu JY, Adler NE, Nguyen R, López A, et al. Social disparities in internet patient portal use in diabetes: evidence that the digital divide extends beyond access. *J Am Med Inform Assoc* 2011 May;18(3):318-321 [FREE Full text] [doi: [10.1136/jamia.2010.006015](https://doi.org/10.1136/jamia.2010.006015)] [Medline: [21262921](https://pubmed.ncbi.nlm.nih.gov/21262921/)]
48. Tieu L, Schillinger D, Sarkar U, Hoskote M, Hahn KJ, Ratanawongsa N, et al. Online patient websites for electronic health record access among vulnerable populations: portals to nowhere? *J Am Med Inform Assoc* 2017 Apr;24(e1):e47-e54. [doi: [10.1093/jamia/ocw098](https://doi.org/10.1093/jamia/ocw098)] [Medline: [27402138](https://pubmed.ncbi.nlm.nih.gov/27402138/)]
49. Showell C. Barriers to the use of personal health records by patients: a structured review. *PeerJ* 2017;5:e3268 [FREE Full text] [doi: [10.7717/peerj.3268](https://doi.org/10.7717/peerj.3268)] [Medline: [28462058](https://pubmed.ncbi.nlm.nih.gov/28462058/)]
50. Logue MD, Effken JA. An exploratory study of the personal health records adoption model in the older adult with chronic illness. *Inform Prim Care* 2012 May;20(3):151-169 [FREE Full text] [doi: [10.14236/jhi.v20i3.21](https://doi.org/10.14236/jhi.v20i3.21)] [Medline: [23710840](https://pubmed.ncbi.nlm.nih.gov/23710840/)]

51. Emani S, Yamin CK, Peters E, Karson AS, Lipsitz SR, Wald JS, et al. Patient perceptions of a personal health record: a test of the diffusion of innovation model. *J Med Internet Res* 2012 Nov;14(6):e150 [FREE Full text] [doi: [10.2196/jmir.2278](https://doi.org/10.2196/jmir.2278)] [Medline: [23128775](https://pubmed.ncbi.nlm.nih.gov/23128775/)]
52. Antonio M, Petrovskaya O, Lau F. Is research on patient portals attuned to health equity? A scoping review. *J Am Med Inform Assoc* 2019 Aug;26(8-9):871-883 [FREE Full text] [doi: [10.1093/jamia/ocz054](https://doi.org/10.1093/jamia/ocz054)] [Medline: [31066893](https://pubmed.ncbi.nlm.nih.gov/31066893/)]
53. Stanford Design School. Equity-Centered Design Framework. URL: <https://dschool.stanford.edu/resources/equity-centered-design-framework> [accessed 2021-07-01]
54. Crawford A, Serhal E. Digital health equity and COVID-19: the innovation curve cannot reinforce the social gradient of health. *J Med Internet Res* 2020 Jun;22(6):e19361 [FREE Full text] [doi: [10.2196/19361](https://doi.org/10.2196/19361)] [Medline: [32452816](https://pubmed.ncbi.nlm.nih.gov/32452816/)]
55. Gordon NP, Lin TY, Rau J, Lo JC. Aggregation of Asian-American subgroups masks meaningful differences in health and health risks among Asian ethnicities: an electronic health record based cohort study. *BMC Public Health* 2019 Nov;19(1):1551 [FREE Full text] [doi: [10.1186/s12889-019-7683-3](https://doi.org/10.1186/s12889-019-7683-3)] [Medline: [31760942](https://pubmed.ncbi.nlm.nih.gov/31760942/)]
56. Islam N, Khan S, Kwon S, Jang D, Ro M, Trinh-Shevrin C. Methodological issues in the collection, analysis, and reporting of granular data in Asian American populations: historical challenges and potential solutions. *J Health Care Poor Underserved* 2010 Dec;21(4):1354-1381. [doi: [10.1353/hpu.2010.0939](https://doi.org/10.1353/hpu.2010.0939)] [Medline: [21099084](https://pubmed.ncbi.nlm.nih.gov/21099084/)]
57. Chin KK. Improving Public Health Surveillance About Asian Americans, Native Hawaiians, and Pacific Islanders. *Am J Public Health* 2017 Jun;107(6):827-828. [doi: [10.2105/AJPH.2017.303802](https://doi.org/10.2105/AJPH.2017.303802)] [Medline: [28498752](https://pubmed.ncbi.nlm.nih.gov/28498752/)]
58. Altman M, Huang TT, Breland JY. Design thinking in health Care. *Prev Chronic Dis* 2018 Sep;15:E117 [FREE Full text] [doi: [10.5888/pcd15.180128](https://doi.org/10.5888/pcd15.180128)] [Medline: [30264690](https://pubmed.ncbi.nlm.nih.gov/30264690/)]
59. Høstgaard AM, Bertelsen P, Nøhr C. Methods to identify, study and understand end-user participation in HIT development. *BMC Med Inform Decis Mak* 2011 Sep;11(1):57 [FREE Full text] [doi: [10.1186/1472-6947-11-57](https://doi.org/10.1186/1472-6947-11-57)] [Medline: [21955493](https://pubmed.ncbi.nlm.nih.gov/21955493/)]

Abbreviations

AAPI: Asian Americans and Pacific Islanders
DISTANCE: Diabetes Study of Northern California
EHR: electronic health record
HIE: health information exchange
HIT: health information technology
PHR: personal health record

Edited by A Kushniruk; submitted 12.02.21; peer-reviewed by N Ratanawongsa, A Garg, E Khoong; comments to author 28.03.21; revised version received 23.05.21; accepted 06.11.21; published 01.04.22.

Please cite as:

Lawrence K, Chong S, Krelle H, Roberts T, Thorpe L, Trinh-Shevrin C, Yi S, Kwon S
Chinese Americans' Use of Patient Portal Systems: Scoping Review
JMIR Hum Factors 2022;9(2):e27924
URL: <https://humanfactors.jmir.org/2022/2/e27924>
doi: [10.2196/27924](https://doi.org/10.2196/27924)
PMID: [35363153](https://pubmed.ncbi.nlm.nih.gov/35363153/)

©Katharine Lawrence, Stella Chong, Holly Krelle, Timothy Roberts, Lorna Thorpe, Chau Trinh-Shevrin, Stella Yi, Simona Kwon. Originally published in *JMIR Human Factors* (<https://humanfactors.jmir.org>), 01.04.2022. This is an open-access article distributed under the terms of the Creative Commons Attribution License (<https://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in *JMIR Human Factors*, is properly cited. The complete bibliographic information, a link to the original publication on <https://humanfactors.jmir.org>, as well as this copyright and license information must be included.

Review

Functional and Technical Aspects of Self-management mHealth Apps: Systematic App Search and Literature Review

Lyan Alwakeel^{1,2}, MSc; Kevin Lano¹, PhD

¹Department of Informatics, King's College London, London, United Kingdom

²College of Computers & Information Technology, University of Tabuk, Tabuk, Saudi Arabia

Corresponding Author:

Lyan Alwakeel, MSc

Department of Informatics, King's College London

Strand

London, WC2R 2LS

United Kingdom

Phone: 44 02078365454

Email: lyan.alwakeel@kcl.ac.uk

Abstract

Background: Although the past decade has witnessed the development of many self-management mobile health (mHealth) apps that enable users to monitor their health and activities independently, there is a general lack of empirical evidence on the functional and technical aspects of self-management mHealth apps from a software engineering perspective.

Objective: This study aims to systematically identify the characteristics and challenges of self-management mHealth apps, focusing on functionalities, design, development, and evaluation methods, as well as to specify the differences and similarities between published research papers and commercial and open-source apps.

Methods: This research was divided into 3 main phases to achieve the expected goal. The first phase involved reviewing peer-reviewed academic research papers from 7 digital libraries, and the second phase involved reviewing and evaluating apps available on Android and iOS app stores using the Mobile Application Rating Scale. Finally, the third phase involved analyzing and evaluating open-source apps from GitHub.

Results: In total, 52 research papers, 42 app store apps, and 24 open-source apps were analyzed, synthesized, and reported. We found that the development of self-management mHealth apps requires significant time, effort, and cost because of their complexity and specific requirements, such as the use of machine learning algorithms, external services, and built-in technologies. In general, self-management mHealth apps are similar in their focus, user interface components, navigation and structure, services and technologies, authentication features, and architecture and patterns. However, they differ in terms of the use of machine learning, processing techniques, key functionalities, inference of machine learning knowledge, logging mechanisms, evaluation techniques, and challenges.

Conclusions: Self-management mHealth apps may offer an essential means of managing users' health, expecting to assist users in continuously monitoring their health and encourage them to adopt healthy habits. However, developing an efficient and intelligent self-management mHealth app with the ability to reduce resource consumption and processing time, as well as increase performance, is still under research and development. In addition, there is a need to find an automated process for evaluating and selecting suitable machine learning algorithms for the self-management of mHealth apps. We believe that these issues can be avoided or significantly reduced by using a model-driven engineering approach with a decision support system to accelerate and ameliorate the development process and quality of self-management mHealth apps.

(*JMIR Hum Factors* 2022;9(2):e29767) doi:[10.2196/29767](https://doi.org/10.2196/29767)

KEYWORDS

mHealth; mobile health apps; mobile apps; apps; systematic literature review; SLR; apps; Mobile App Rating Scale; MARS; smartphone; iOS; Android; mobile phone

Introduction

Self-management mobile health (mHealth) apps use mobile devices for health services and offer a sustainable means of enhancing self-health care management to achieve wellness goals, such as health monitoring, disease detection, behavior change, and emotion management, enabling individuals to independently manage their lives and activities and make appropriate decisions. They comprise different categories, ranging from simple apps for water intake tracking to complex apps that can adapt to individuals' lives based on their activities. The power of self-management mHealth apps has increased with the use of built-in mobile technologies (eg, cameras, GPS, and accelerometers) and machine learning (ML) algorithms to create intelligent mobile apps. Such apps are characterized by personalized services and recommendations and by the automatic logging and recognition of individuals' behaviors and activities.

Although mobile apps are extensively used for self-health care management, the ongoing development of mobile device technologies and programming languages has increased the need for mobile app solutions to keep pace with development practices. Developing high-quality intelligent self-management mHealth apps requires substantial knowledge of mobile programming languages, app architectures, design patterns, and latest technologies. Considerable time is required from researchers and developers to learn and master such knowledge because of the different characteristics, requirements, and components of each mobile app. Furthermore, many challenges and issues may arise during app development. Therefore, we conducted a comprehensive systematic literature review (SLR) and evaluation of existing self-management mHealth apps that focus on self-health care management, based on a formal protocol, to analyze their characteristics and current challenges, including infrastructure, functionalities, user interface (UI) components, screen navigation, services and technologies, security and authentication, use of architectures and patterns, evaluation, and issues to provide a guide for self-management mHealth app infrastructure to facilitate further development.

Although several SLRs on mHealth apps have been previously conducted, our SLR is distinguished by providing engineering perspectives on 3 different sources: research papers, app stores, and GitHub repositories. The SLR involved the analysis and synthesis of empirical evidence by software engineering

researchers to help researchers and developers in three main aspects: (1) to identify the characteristics and challenges of existing self-management mHealth apps, (2) to understand the differences and similarities of existing self-management mHealth apps and find the gap between research papers and commercial and open-source apps, and (3) to suggest future research directions based on gaps identified in the domain. The main contributions of this study are as follows:

- The definition of an SLR protocol following the SLR guidelines by Kitchenham and Charters [1], which is based on a wide range of literature and the selection of 52 research papers, 42 app store apps, and 24 open-source apps as primary studies
- Extraction, analysis, synthesis, and reporting of empirical evidence from the selected primary studies
- Provision of guidance for researchers and developers to deeply understand the characteristics and challenges of self-management mHealth apps
- Suggestions of solutions to overcome the limitations of existing self-management mHealth apps

The principal aim of this SLR was to obtain a detailed view of existing self-management mHealth apps used for self-health care management. A specific objective was to better characterize the functional and technical aspects of these apps.

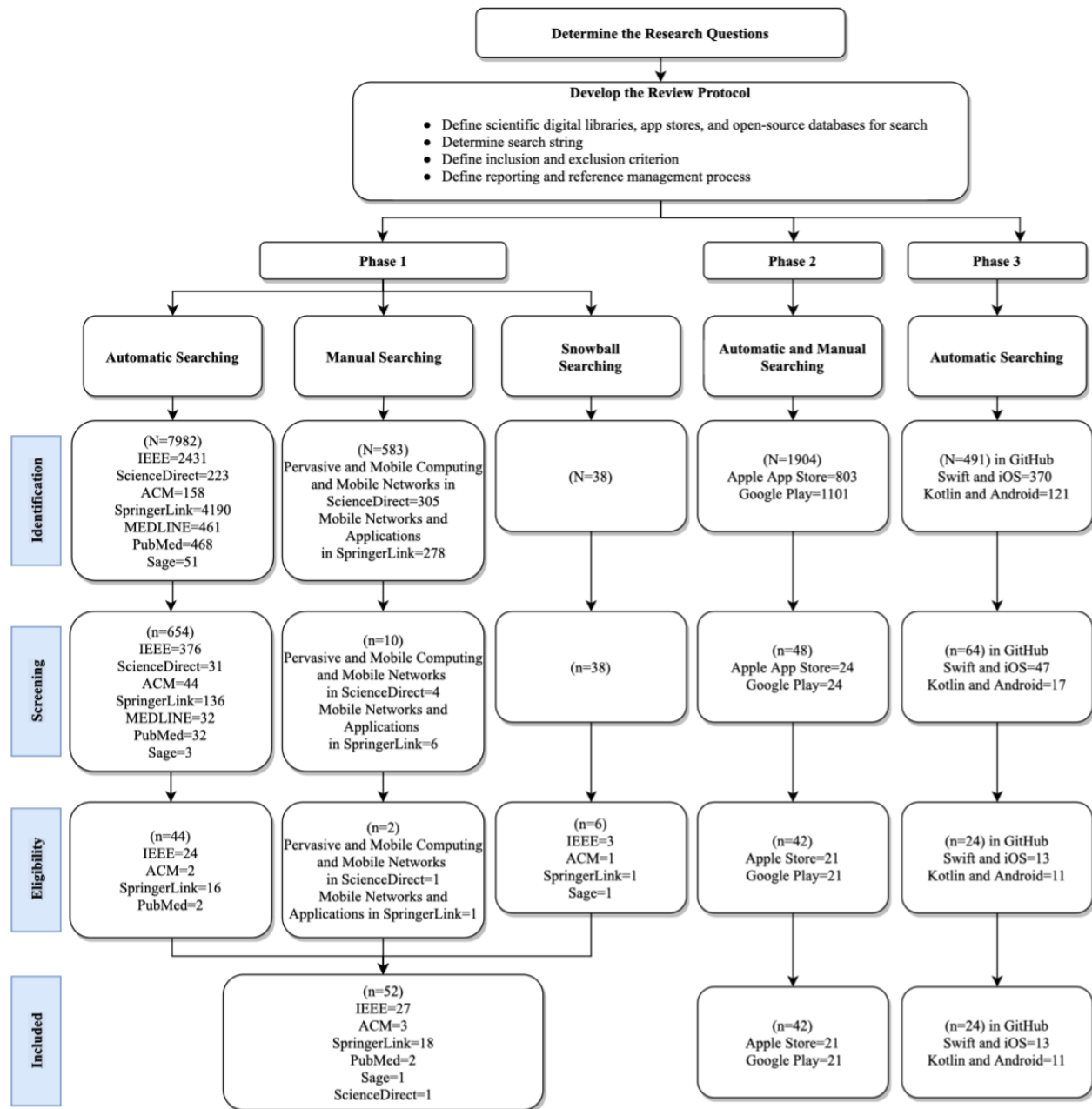
Methods

Study Design

This review presents the main characteristics and challenges of the self-management of mHealth apps. We targeted apps that use mobile devices for the self-health care management of general users. The study process diagram is shown in [Figure 1](#) where it is divided into 3 main phases for analyzing apps that exist in the knowledge base:

1. Phase 1: a comprehensive review of existing research papers on self-management mHealth apps in digital libraries
2. Phase 2: an exploration of self-management mHealth apps available through Britain's Apple App Store and Android Google Play, as well as an evaluation of the selected apps using the Mobile App Rating Scale (MARS) [2]
3. Phase 3: analysis of open-source apps available on GitHub based on specific criteria and an automatic tool (ie, SonarCloud; SonarSource SA) [3]

Figure 1. Process diagram for this systematic literature review. mHealth: mobile health.



Review Methodology

This section presents related literature in the knowledge base, which forms the foundation of our SLR. The SLR follows the Kitchenham and Charters [1] guidelines, which divide the review of each phase into 3 main stages of the review: planning, conducting, and reporting. All the stages were prepared by the first author and revised by the second author. The following subsections outline the steps that were followed.

Research Questions

Overview

To investigate and deduce empirical evidence of existing self-management mHealth apps, we determined 2 key research questions (RQs):

- RQ1: What are the main characteristics of current self-management mHealth apps?
- RQ2: What are the challenges and issues faced by current self-management mHealth apps?

Detailed Explanation of the RQs

The Characteristics of Self-management mHealth Apps

We determined some characteristics of self-management mHealth apps that we were interested in monitoring because of their importance in the development of mobile apps. We collected these characteristics based on their availability during each phase (Textbox 1).

Textbox 1. Characteristics of self-management mobile health apps.

Crucial functionalities

- Each app comprises several components that define its functionality. We have summarized the main functionalities of the reviewed apps.

User interface components

- We determined the user interface components used by the users for interaction.

Navigation and structure

- Here, we explored the apps' organization and methods of navigating the app screens.

Services and technologies

- We determined the remote and local services that are external to the assigned app but handled by it, such as machine learning algorithms; built-in technologies; and access to other apps, frameworks, and libraries.

Security features

- We were interested in defining the security aspects and authentication mechanisms used in the self-management of mobile health (mHealth) apps.

Architectures and patterns

- There are different architectures and patterns for building apps, such as client-server, model-view-controller, and view-interactor-presenter-entity-router. In this study, we determined the most commonly used architectures and patterns. Furthermore, we specified the architecture for implementing machine learning, including web-based inference, offline inference, or both.

Logging mechanisms

- We determined the apps' method for logging data, either manually or automatically.

Development approach

- We explored the main development approaches which developers use to construct apps.

Operating system and programming language

- We identified the operating system and programming languages used in the reviewed apps.

Evaluation

- Here, we were concerned with the techniques for evaluating self-management mHealth apps used by researchers in phase 1. Furthermore, we evaluated the selected apps using Mobile App Rating Scale [2] and SonarCloud (SonarSource SA) tools to assess their quality in phases 2 and 3. We used the Mobile App Rating Scale in phase 2, which is a reliable tool developed by an expert panel for evaluating the quality of mHealth apps and comprises an initial section for gathering general and technical information about the assigned app and 5 specific sections: engagement, functionality, aesthetics, information quality, and subjective quality. Each section has a group of items that can be scored from 1 (inadequate) to 5 (excellent). These scores are used to calculate the mean score for each section. Finally, the average values of the mean of the first 4 sections (ie, engagement, functionality, aesthetics, and information quality) are calculated to obtain the final measurements of app quality. All apps were evaluated and compared to find the differences between platform versions. In phase 3, we used a web-based service (SonarCloud) for static code analysis, as well as manual exploration to identify app characteristics.

The Challenges and Issues of mHealth Apps

We identified the limitations and challenges faced by researchers and developers when developing self-management mHealth apps. In phase 1, we summarize the researcher's challenges and issues. In phases 2 and 3, we identified potential issues that can affect the quality of apps using MARS and SonarCloud, such as design issues, bugs, code smell, and duplication.

Search Strategy

Overview

A search strategy was followed to explore the literature that could help answer the RQs. It comprises 3 main stages: defining search strings, selecting data sources, and searching the data

sources. As previously mentioned, the review is divided into 3 independent phases, each of which has separate data sources and search strings. In the first phase, we followed the quasi-gold standard [4] approach, including manual and automatic searches, as well as snowballing. The second phase included manual and automatic searches, whereas the third phase was limited to automatic searches.

Defining Search Strings

The search string of the first phase was defined by combining synonymous terms using *OR* and *AND*. On the basis of our RQs, 5 search strings were identified, as listed in [Textbox 2](#).

In the second phase, an extensive search was performed in Britain's Apple Store and Google Play from August 16 to

August 21, 2020. Although the search was limited to Britain's app stores because of the requirement to specify the user's location, most of the selected apps were available in other stores. We started the automatic search by applying the following search string: *mHealth*, *Healthcare*, and *Health*. However, many unrelated apps were identified. Consequently, we changed the search strategy to a manual exploration of the *Top Free App*

under the *Health and Fitness* category. These apps are free to download, although many require a monthly payment or upgrade payment to access all features.

In the third phase, we used the search strings *mHealth*, *Healthcare*, and *Health* in GitHub from October 5 to October 9, 2020, to identify open-source mHealth apps.

Textbox 2. Search string used for digital libraries.

Search strings

1. "mHealth" AND ("app" OR "application").
2. "mobile" AND "health" AND ("app" OR "application").
3. "Personal" AND "Mobile" AND "healthcare" AND ("app" OR "application").
4. "Self-management" AND "healthcare" AND ("app" OR "application").
5. "Smartphone" AND "health" AND ("app" OR "application").

Data Sources

To find studies related to phase 1, we followed an automatic search using the following digital libraries: IEEE Xplore, ScienceDirect, ACM Digital library, SpringerLink, MEDLINE, PubMed, and Sage. To complement the automatic search, a manual search was conducted on relevant journals, including Pervasive and Mobile Computing and Mobile Networks and Mobile Networks and Applications. To find as many studies as possible, we used the snowballing strategy to gather additional studies from the reviewed studies. We used Google Scholar to search for further studies identified from snowballing. In phase 2, we used an automatic and manual search within the following official digital British app stores: Apple iPhone (App Store) and Android (Google Play). In the third phase, we applied an automatic search on GitHub to download open-source apps.

Search Process in Data Sources

To identify all related studies, search strings were applied to the selected digital libraries. Initially, 7982 results were retrieved within the chosen search string as follows: 2431 (30.46%) papers from IEEE, 223 (2.79%) papers from ScienceDirect, 158 (1.98%) papers from ACM, 4190 (52.49%) papers from SpringerLink, 461 (5.78%) papers from MEDLINE, 468 (5.86%) papers from PubMed, and 51 (0.64%) papers from Sage. The results were filtered based on the title and abstract, and duplicated and unrelated papers were removed. Then, of the 7982 papers, 654 (8.19%) were downloaded for examination. According to the inclusion and exclusion criteria, of the 654 studies, 44 (6.7%) studies were included from the automatic search. Furthermore, we found 2 studies from the manual search and 6 studies from snowballing. Thus, we collected 52 studies from automatic and manual searches, as well as snowballing. Manual search studies were from ScienceDirect and SpringerLink. For the 6 studies from snowballing, 3 (50%)

papers were from IEEE, and 1 (17%) paper each was from SAGE, ACM, and SpringerLink.

The systematic search process was applied to identify mobile apps for the general population in phase 2. Initially, 1904 apps were retrieved with the chosen search strategy: 803 (42.17%) apps from Apple's App Store and 1101 (57.83%) apps from Android's Google Play. The results were filtered based on the inclusion and exclusion criteria and availability in both app stores. After removing duplicate apps, we downloaded and explored 100% (48/48) of apps on each platform, iOS and Android. A total of 48 apps were reviewed for further refinement after applying the inclusion and exclusion criteria. Then, 6% (3/48) of apps from each store were removed as they were not used for self-health care management. Therefore, of the 42 apps, the final number of apps included on each platform was 21 (50%).

In phase 3, the systematic search process was applied to GitHub to identify self-management mHealth apps for the general population using the Swift and Kotlin programming languages. Initially, 491 apps were retrieved with the chosen search strategy as follows: 370 (75.4%) apps with Swift programming language and 121 (24.6%) apps with Kotlin language. After removing duplicates and unrelated titles, of the 491 apps, we obtained 64 (13%). These apps were downloaded and analyzed based on the inclusion and exclusion criteria. After applying the inclusion and exclusion criteria, of the 64 apps, we obtained 24 (38%). These apps were analyzed using the SonarCloud tool, including 13 (54%) iOS and 11 (46%) Android apps.

Inclusion and Exclusion Criteria

Textbox 3 presents the inclusion and exclusion criteria that were applied to the downloaded papers that resulted from the manual, automatic, and snowballing search of digital libraries, as well as apps that resulted from digital app stores and GitHub.

Textbox 3. The inclusion and exclusion criteria of self-management mobile health apps.

Phase 1

- Inclusion criteria
- Papers presenting the design and implementation of either or both Android and iOS self-health care management apps
- English peer-reviewed papers published from 2008, the year the App Store was announced [5], to 2020
- The most recent and complete version of a study if it had multiple versions
- Exclusion criteria
- Papers presenting theoretical research without implementation
- Papers describing apps for wearable devices
- Papers targeting children or people with special needs
- Short papers with <4 pages as they could not contain sufficient information

Phase 2

- Inclusion criteria
- Apps designed for general users, existing in both Apple and Android app stores
- Apps that were free to download and could support the English language
- Apps stating the aim as self-health care management
- Apps rated by >10,000 users and having a score of ≥ 4 out of 5 to ensure that the selected apps were satisfying
- The last update of the current app had to be from January 01, 2018 to ensure that the app was up to date
- Exclusion criteria
- Apps designed for children or people with special needs
- Apps related to an external device such as a smartwatches or shoes
- Apps that did not clarify the date of the last update

Phase 3

- Inclusion criteria
- Apps written using either or both Swift or Kotlin
- Open-source apps that supported the English language
- Apps stating the aim as self-health care management
- The source code of the app had to exceed 1000 lines of code
- Exclusion criteria
- Apps designed for children or people with special needs
- Apps related to external devices such as smartwatches or shoes

Overview of the Selected Studies

In phase 1, of 52 studies, 27 (52%) were from IEEE, 18 (34%) from SpringerLink, 3 (6%) from ACM, 2 (4%) from PubMed, 1 (2%) from ScienceDirect, and 1 (2%) from Sage. MEDLINE digital data sources were not included as the downloaded papers focused on analyzing user behavior rather than app development, which was outside the scope of this study. In phase 2, we reviewed 21 apps on each platform, each of which had 2 versions: 1 in the Apple App Store and 1 in Android Google Play. In phase 3, a group of 24 open-source apps was reviewed manually and automatically, including 11 (46%) Android and 13 (54%) iOS apps.

In total, we reviewed 41 iOS and 77 Android apps, including 7 (17%) iOS and 45 (58%) Android apps in phase 1, 21 (51%) iOS and 21 (27%) Android apps in phase 2, and 13 (54%) iOS and 11 (46%) Android apps in phase 3. The apps in phase 2 have almost the same functionalities and structures on both platforms. Thus, we use a letter with a number (eg, *A1*) to represent the app name of the 2 versions, and we specify the differences if they are found on each app.

Results

Overview

Here, we summarize the findings obtained by reviewing the selected papers and apps based on our RQs. The general context of apps from the data extracted in phases 1, 2, and 3 are

presented in [Multimedia Appendix 1](#) [6-81]. The main characteristics of the 3 phases are presented in [Multimedia Appendix 2](#).

The Characteristics of Self-management mHealth Apps

Crucial Functionalities

Self-management mHealth apps had several focuses, including physical health, weight control, sleep, mental health, disease, women’s health, and monitoring, as shown in [Figure 2](#), where physical health and weight control were the most frequent focus in the 3 phases. These apps used various terminologies to describe their crucial functionalities, including detection, recognition, prediction, estimation, monitoring, personalization, and recommendations. The word *detection* was used to detect whether there was something abnormal in data, such as a disease, whereas the word *recognition* was used to recognize the type of something, such as the type of specific activity or food. Prediction and estimation can use collected data to predict

a situation or estimate a value. Monitoring depends on the calculation methods used to monitor the user’s progress. Regarding personalization and recommendation, the apps provided customized plans, guidelines, and suggestions based on the users’ data and their progress. [Figure 3](#) shows the frequency of studies and apps, including their functionalities in the 3 phases. Specifically, 60% (31/52) of the studies in phase 1 focused on recognition. However, 81% (34/42) of the apps in phase 2 and 83% (20/24) of the apps in phase 3 focused on monitoring. In contrast, a few studies in phase 1 demonstrated the usual app functionalities such as log-in and analysis as they focused on presenting their new contributions in developing ML algorithms. Phases 2 and 3 profusely included usual functionalities, including log-in, payment, synchronization of data from other apps, rating or questionnaire, search, sharing data using email, WhatsApp, and Telegram, and analysis to periodically visualize reports or charts to help users easily read results and achieve their desired goals.

Figure 2. The general focus of reviewed apps.

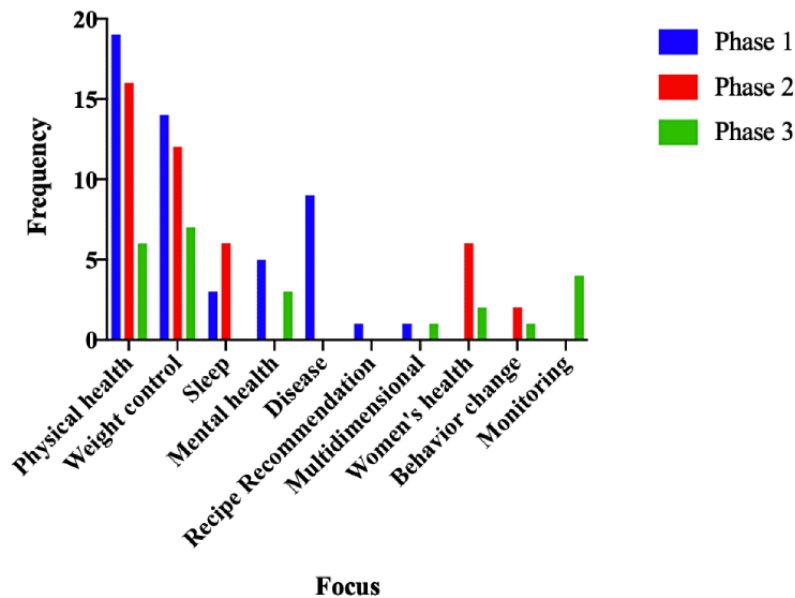
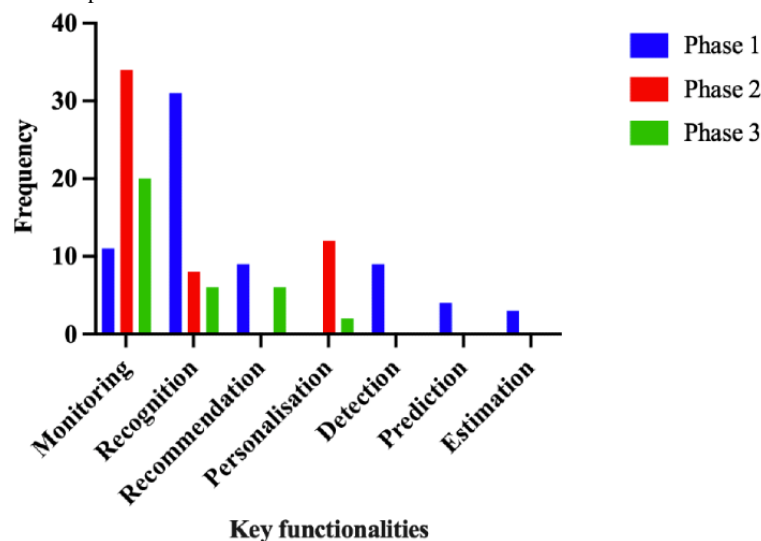


Figure 3. The key functionalities in the 3 phases.



UI Components

The most frequent UI components in the 3 phases were labels, images, buttons, input boxes, lists, visualized charts to present user progress, menu or hamburger menu, and radio buttons. Other components such as tables, visual control bars, and virtual pets appeared less frequently. Furthermore, apps in phases 2 and 3 had more frequency of the following UI components than phase 1: switch, scroll bar, picker, calendar, and media.

Navigation and Structure

To navigate through an app's screens, the next and back buttons, tab structure, hamburger menu, and home page could be used. In phase 1, the tab structure appeared in 12% (6/52) of studies. The next and back buttons were used in 10% (5/52) of studies. One of the studies included navigation through both the next and back button and the tab, whereas another study supported navigation with the next and back button, tab pages, and hamburger menu. Approximately 4% (2/52) of studies reported that the user could use the home page to navigate to other pages.

Most apps (40/42, 95%) in phase 2 had, at start-up, multiple simple pages for customization with backward and forward arrows; however, the basic app was presented as a tab structure with multiple tabs. In the case of iOS, of the 42 apps, 18 (43%) were designed with a tab structure, 2 (5%) supported navigation with tabs and a hamburger menu, and the remaining 1 (2%) app had a main page with hamburger menu. For Android, of the 42 apps, 17 (40%) followed the tab structure, 3 (7%) had a page with a hamburger menu, and 1 (2%) had a main page with buttons to navigate to other pages. Consequently, 86% (36/42) of apps had a similar structure on both platforms, where 81% (34/42) of the apps followed the tabbed UI approach, and approximately 5% (2/42) of the apps had a main page with hamburger menu.

In phase 3, 71% (17/24) of apps had a tabbed UI structure, including 33% (8/24) having a tab with a back arrow to navigate to the previous or main page, 29% (7/24) of apps supporting navigation through only tabs, and 8% (2/24) of apps supporting navigation through both tabs and hamburger menu. Approximately 29% (7/24) of apps had a main page structure, including 8% (2/24) supporting a main page with a back arrow, 8% (2/24) having a main page with a back arrow and menu, 8% (2/24) having a main page with navigation buttons to other pages, and 4% (1/24) having a main page with a menu.

Services and Technologies

Several local and remote services were used in the reviewed studies and apps. These services included accessing remote and local libraries and technologies; accessing external libraries using an application programming interface (API) such as Clarifai, ZXing (Zebra Crossing), Edamam, Dialogflow, and ToneAnalyzerV3 (IBM Watson); web services; cloud services such as Google's cloud computing engine (Firebase); linking to other apps such as social media; and Google Maps. In addition, some apps used developer frameworks such as HealthKit, ARKit, SceneKit, and StepCounter.

Many studies and apps in the 3 phases had access to built-in technologies such as cameras, motion sensors (accelerometers

and gyroscopes), location (GPS), and microphones. Phases 2 and 3 had access to photographs, vibrations, networks, audio, phones, and storage. These technologies can collect various types of data, including sensor data, images, voice, and text. The collected data can be used to process through ML algorithms to obtain useful information.

Several studies in phase 1 used ML algorithms that differentiated between supervised and unsupervised learning with their branches: classification, regression, clustering, and association. Some studies evaluated multiple algorithms to determine the most accurate algorithm. Most of the investigated studies used supervised learning for classification, including naïve Bayes, support vector machines, logistic regression, K-nearest neighbor, rule-based classifiers, decision trees, ridge, AdaBoost, bagging, Gaussian processes, ensemble of nested dichotomies, rotation forest, Fisher vector representation, linear classifiers, and artificial neural networks (NNs), which included specific types such as deep NNs and deep convolutional NNs. Moreover, supervised learning was used for regression, which comprised linear regression (LR), Bayesian ridge, support vector regression, gradient boosting, and AdaBoost. Some studies used algorithms for both classification and regression, such as random forest. Other studies used unsupervised learning for clustering, such as the density-based spatial clustering of apps with noise and molecular complex detection. Other used algorithms were the kernel null Foley-Sammon transform and t-distributed stochastic neighbor embedding, the threshold method, decision tables, radial basis function kernel, ensemble extreme learning machine, and sequential minimal optimization. In contrast, many studies used measurement methods, including BMI, basal metabolic rate, Gaussian and LR functions, and general calculations.

Security Features

Authentication through log-in was the principal security feature in the 3 phases, which was achieved either by creating an app account using email or linking it with other accounts such as Facebook, Apple, or Google. However, only one of the studies (S16) mentioned authentication through log-in in phase 1. This limitation could be because of the focus on building and evaluating a new ML algorithm rather than a complete app. Approximately 95% (40/42) of apps in phase 2 and 42% (10/24) of apps in phase 3 supported log-in. The log-in password was hidden using points or stars.

Architectures and Patterns

In phase 1, some studies involved the development of mobile apps to collect real data; however, they analyzed these data on a physical computer, server, or specific tool (eg, Waikato Environment for Knowledge Analysis or MATLAB) to extract features and identify the most accurate ML algorithm. Most of these studies focused on building a suitable model using ML algorithms without providing real information about the mHealth app, which integrates the final model with implementation or architecture. These apps stored the automatically-collected data remotely on a server (2/52, 4%) using a client-server architecture (a web-based mechanism) or locally (7/52, 13%) on a mobile device (an offline mechanism). Approximately 2% (1/52) of studies used pre-existing data sets for processing on a physical

computer to build the model. Other studies focused on developing complete self-management mHealth apps, in which 54% (28/52) of studies used a web-based mechanism for developing the apps. They followed a client-server architecture, where a mobile app acts as a client node to receive data and connect to a server (sometimes a cloud server) to process data or perform computations. Other studies (12/52, 23%) used an offline mechanism to process received data locally on mobile devices. One of the studies (S24) in phase 1 followed MobileNetV2 (deep NN architecture). In S13, the authors used cloud-based virtualization (hypervisor architecture), which depends on virtual swaps between mobile sessions to separate the physical resources into virtual resources on a cloud server to use computing power efficiently. Consequently, in phase 1, 58% (30/52) of the studies used a web-based mechanism, and 37% (19/52) used an offline mechanism. In phase 3, 83% (20/24) were web-based apps and 17% (4/24) were offline apps.

Regarding the design pattern, 2 authors (S38 and S40) in phase 1 stated that they used the model-view-controller (MVC) design pattern, whereas, in phase 3, a total of 3 design patterns were used in the explored apps: MVC, model-view-viewmodel, and view-interactor-presenter-entity-router. MVC was used in 75% (18/24) and model-view-viewmodel in 21% (5/24) of the apps. View-interactor-presenter-entity-router was used in 4% (1/24) of apps.

In terms of the architecture for implementing ML mobile apps in phase 1, of the 52 studies, 18 (35%) used web-based inference based on a pretraining model by researchers, 10 (19%) used offline inference on devices based on a pretraining model by researchers, 3 (6%) used both web-based and offline inference based on a pretraining model by researchers, and 2 (4%) used web-based inference based on ready solutions (API), whereas, 29% of ML apps in phase 3 used web-based inference based on ready solutions (API). Web-based and offline inferences were used for the same functionality in S30 and S51, where the web was used to accelerate the process, and offline was used when a connection was lost but with lower performance and high resource consumption. However, S1 used both web-based and offline inferences for different functionalities.

Logging Mechanism

Most studies and apps collected data from user information, activities, or behaviors, which could be gathered automatically from sensors, manually by users' logging, or both. Automatic collection can be achieved either through synchronization and importing of data from other apps or by accessing built-in technologies or developer frameworks such as HealthKit on iOS or Google Fit on Android. Then, the collected data were automatically analyzed using ML algorithms or calculation methods to provide useful feedback and personalized services.

In phase 1, most studies (24/52, 46%) used built-in technologies for automatic data collection. Other studies required manual input of information (8/52, 15%), image capture (14/52, 27%), or voice recording (2/52, 4%). Few studies (4/52, 8%) supported both automatic collection and manual inputs.

In phase 2, manual logging was used in 62% (26/42) of apps and could be performed through barcode scanning, input of

personal information, or searching the app's internal database. Approximately 24% (10/42) of the apps supported automatic logging through built-in technologies or synchronization with other apps. Approximately 14% (6/42) of the apps supported both automatic and manual logging.

Most apps (18/24, 75%) in phase 3 supported manual logging of data such as personal information, food and water consumption, sleep, and emotion. Approximately 17% (4/24) of the apps supported manual logging and synchronization. One of the apps presented real-time information from an external server, and another app supported the automatic collection of data.

Development Approach

Prototyping was the most commonly used approach adopted in 15% (8/52) of studies. Other studies used agile (S35), extreme programming (S15), iterative development (S38), or user-centered approaches (S26 and S29). The remaining studies in phase 1 and the other phases did not mention their approach.

Operating System and Programming Language

In phase 1, 85% (44/52) of studies targeted Android, forming the majority. Approximately 13% (7/52) supported iOS, and 2% (1/52) targeted both iOS and Android. In terms of operating system (OS) versions, S2 used Android OS version 4.4.2; S6 and S28 used Android OS version 4.1.2; S22 used Android OS version 1.6; S29 used Android OS version 4.2.2 (Jellybean); S41 used Android OS version 2.3; and S39 targeted both platforms, with Android version 4.0 (Ice Cream Sandwich) and iOS version 3.2. The Java programming language was used in S1, S3, S16, S17, S21, S28, S31, S34, S40, and S38. Plain Old Java Objects were used in S42. Each selected app in phase 2 was available on both the Android and iOS OSs. We reviewed and evaluated 24 apps in phase 3, 11 (46%) of which had Android OS and 13 (54%) of which had iOS.

Evaluation

Different evaluation techniques were used in phase 1; however, most studies (38/52, 73%) measured the performance of ML algorithms through experiments, pilot studies, or randomized controlled trials and compared the performance with state of the art. Furthermore, some studies used a specific tool to test the accuracy of various classifiers and select the most appropriate one. For example, 12% (6/52) of the studies used the Waikato Environment for Knowledge Analysis, and approximately 6% (3/52) used MATLAB. Furthermore, cross-validation was used in approximately 17% (9/52) of the studies to accurately calculate the performance metrics. Such metrics could include confusion matrix, sensitivity, specificity, and accuracy. Other studies (4/52, 8%) evaluated through a usability study, such as user acceptability and subjective surveys. Approximately 10% (5/52) of the studies combined either or both comparisons with state-of-the-art and usability studies.

In phase 2, we applied a systematic quality evaluation of the apps selected from app stores. The evaluation was conducted by the first author and revised by the second author. Each app was opened on both the Android and iOS platforms for evaluation. We evaluated the apps using MARS, comprising 5

main sections (engagement, functionality, aesthetics, information, and subjective quality), which are presented in [Multimedia Appendix 3](#), where the tables list the scores for each section and the final mean score of every app. The app's section score was calculated by taking the average score of each item for each app. A2 and A17 received the highest score (4.6) in the engagement section, whereas A10 and A16 received the lowest score (3.2). The mode of the functionality section was 4.25, whereas it was 4.7 in the aesthetics section. The information section had the highest score (4.9) in 2 apps (A3 and A15), whereas the other apps received a score of 4.7.

[Table 1](#) reports the mean scores for an overall score of the quality of each app and the subjective quality. The overall score

of the app quality was the average of the section scores (excluding subjective quality, which was calculated separately). The median overall mean score was 4.46/5. A2 and A17 received the highest overall score (4.56/5), whereas A10 had the lowest score (3.83/5). In subjective quality evaluation, A6 and A17 received the highest score (4.25/5), and A13 had the lowest score (2/5).

In phase 3, we evaluated the apps using SonarCloud, as shown in [Multimedia Appendix 4](#). We found that most apps had a relatively small number of lines of code, ranging from 1.1 to 7 K.

Table 1. Overall and subjective Mobile App Rating Scale evaluation of self-management mobile health apps.

App ID ^a	Overall score, mean (SD)	Subjective quality, mean (SD)
A1	4.46 (0.27)	3.75 (0.50)
A2	4.56 (0.21)	4 (0.82)
A3	4.06 (0.60)	2.75 (1.26)
A4	4.45 (0.33)	2.75 (1.26)
A5	4.51 (0.22)	4 (0.82)
A6	4.38 (0.44)	4.25 (0.96)
A7	4.14 (0.47)	2.25 (0.96)
A8	4.46 (0.27)	3.25 (1.71)
A9	4.19 (0.51)	2.75 (1.26)
A10	3.83 (0.64)	2.25 (0.96)
A11	4.14 (0.47)	2.25 (0.96)
A12	4.46 (0.27)	3.25 (1.71)
A13	4.45 (0.33)	2 (1.41)
A14	4.14 (0.47)	2.25 (0.96)
A15	4.38 (4.45)	3.25 (1.71)
A16	4.04 (0.63)	3 (1.41)
A17	4.56 (0.21)	4.25 (0.96)
A18	4.32 (0.30)	2.5 (1.0)
A19	4.51 (0.22)	4 (0.82)
A20	4.36 (0.43)	2.75 (1.26)
A21	4.51 (0.22)	2.75 (1.26)

^aApp ID represents the app name of the 2 versions, and we specify the differences if they were found for each app.

The Challenges and Issues of Self-management mHealth Apps

Some studies in phase 1 mentioned general challenges related to mobile devices and app architecture. The first challenge was the restricted number of resources (10/52, 19%), including computational power, storage capacity, and energy efficiency. Other studies referred to the challenge of dealing with a large variety of mobile devices with different software and hardware, which complicates the development of new algorithms (S28) and causes varying levels of accuracy when data are collected

from sensors (S32). Furthermore, S17 mentioned security as a challenge for mobile apps.

In contrast, S23 summarized the drawbacks of cloud-based approaches (a web-based mechanism): latency, privacy, cost, and connectivity. S23 also mentioned the limitations of the offline mechanism that integrates the model in the app, which requires republishing a new version of the app with each update of the model and could be inconvenient for the user and result in a waste of time. S42 mentioned the challenge of designing mHealth apps that succeeded in attracting and sustaining users' interests.

Other studies encountered difficulties in collecting and dealing with the collected data, such as identifying the position of the device on a user's body when collecting data in 8% (4/52) of studies; determining the set of collected sensors in 8% (4/52) of studies; extracting efficient features from collected data in 8% (4/52) of studies; accurate recognition of activities in real time in 6% (3/52) of studies; and accurate detection of heart rate in 4% (2/52) of studies, which depends on the lighting conditions and location of the finger on the camera lens.

In terms of image processing, the authors of S45 mentioned the limitations of mobile devices in dealing with complex images for extracting features and classifying algorithms. The authors of S23 and S15 mentioned that the properties of an image might be affected by various factors, such as the angle, brightness, focal distance, and camera resolution. The authors of S15 and S49 stated that food recognition could be difficult because of several factors, such as diversity of food size, form, color, and texture, as well as deformation and segmentation of food elements, which may affect the identification of food type and calculation of quantity and nutritional value.

In phase 2, we found that most apps (26/42, 62%) used manual logging of data, such as manually inputting the type and duration of the exercise, user emotion, and the category and quantity of consumed food, without using ML algorithms such as automatic recognition of activity and food type. In addition, all explored apps targeted general users without experimental or clinical studies to support their safety, reliability, and effectiveness. We also noticed that most apps (26/42, 62%) needed the internet to access some functions, two of which supported offline working through downloading of content. Approximately 33% (14/42) of apps could work without an internet connection; and 5% (2/42) of apps could not work at all without the internet. We also checked whether the installed apps were hybrid or native by activating the developer option on the Android OS from settings and by turning on the layout bounds option. We found that most Android apps (18/21, 86%) were native. However, this method was not applicable to apps developed using react-native or flutter as they convert the language to native app code. Therefore, the number of native apps was not completely accurate. However, it clarifies to some extent that separate implementations need to be written for each platform, which requires additional time, cost, and effort.

Each app in phase 3 was analyzed and evaluated using SonarCloud, as shown in [Multimedia Appendix 4](#). Of the 24 apps, bugs existed in 7 (29%)—1 (4%) in Android and 6 (25%) in iOS. Approximately 17% (4/24) of apps on iOS had vulnerability issues. Of the 24 apps, 23 (96%) on both platforms had code smells. The highest number of code smells was 181, whereas the lowest was 3. Most apps (18/24, 75%) had duplications, ranging from 0.3% to 14%. Furthermore, we tried to run the apps, and of the 24 apps, only 11 (46%) worked, including 7 (29%) Android and 4 (17%) iOS.

Discussion

Comparison and Synthesis of Phases

In this section, we compare and synthesize the collected data and findings from the first, second, and third phases to answer our RQs.

By comparing the characteristics and challenges of the 3 phases, we identified that the phases were different in terms of the use of ML, processing techniques, functionalities, inference of ML, logging mechanisms, evaluation techniques, and challenges. However, they were similar in the most frequent focus, UI components, navigation and structure, services and technologies, authentication features, and architecture and patterns. As shown in [Tables 2, 3, and 4](#), we reviewed 52 studies in phase 1, 21 apps in phase 2 (each of which has 2 versions), and 24 apps in phase 3. We found that most studies of phase 1 (43/52, 83%) were intelligent and used ML algorithms, supporting supervised learning (39/52, 75%), unsupervised learning (1/52, 2%), both supervised and unsupervised learning (1/52, 2%), and accessing external ML libraries through API (2/52, 4%). Most supervised learning studies (34/52, 65%) focused on classification. Some studies used ML in phase 2 (14/42, 33%) and phase 3 (7/24, 29%), where phase 3 depended on an external ML library (API).

In terms of processing techniques, most studies from phase 1 used data (25/52, 48%) and image (11/52, 21%) processing, whereas calculation methods were the most used techniques in phase 2 (22/42, 52%) and phase 3 (13/24, 54%). Data processing included sensor data, questionnaire answers, conversation, and specific data such as goals and preferred meals. The calculation depended on specific equations such as BMI. With respect to focus, the 3 phases were similar, with the most frequent focuses being on physical health and weight control. However, they differed for crucial functionalities, where recognition (20/52, 38%) and detection (9/52, 17%) were the most frequent functionalities in phase 1. Monitoring was the most crucial function in phases 2 and 3, representing 52% (22/42) and 54% (13/24), respectively. The 3 phases were almost similar in terms of UI components and navigation. The most commonly used UI components were labels, images, and buttons, and most apps were designed with a tab structure.

Regarding services and technologies, the camera, GPS, motion sensors, and microphones were frequently used in the 3 phases. The motion included access to accelerometers and gyroscopes and had the highest percentage (21/52, 40%) in phase 1, whereas the camera (20/42, 48%) was the most frequent built-in technology in phase 2, and GPS (7/24, 29%) had the highest percentage in phase 3. Apps in all phases used log-in functionality as a security feature for authentication, which could be achieved by creating a new account with the app using a social media account. For architecture and patterns, most studies in phase 1 (30/52, 58%) and apps in phase 3 (20/24, 83%) used client servers. MVC was the only pattern used in phase 1 and the most used pattern in phase 3 (18/24, 75%). However, phase 2 did not contain an architecture section as there was insufficient information about it in the app stores. In terms of inference of ML, many apps (18/52, 35%) in phase 1 used web-based inference using a pretrained model by

researchers. Phase 3 concentrated on web-based inference, with 29% (7/24) of apps using ready solutions, such as the IBM Watson API. Furthermore, the prototype was the most commonly used development approach in phase 1. However, phases 2 and 3 did not include the development approach section because of limited information.

Most apps (46%) in phase 1 were distinguished by automating the logging mechanism using built-in technologies and ML algorithms to automatically recognize types, quantities, and calories of food and physical activities such as walking, running,

or jumping with burned calories. Apps in phases 2 and 3 (26/42, 62%, and 18/24, 75%, respectively) concentrated on monitoring functionality through manual logging of specific activities such as eating an apple or walking.

We used different techniques for the evaluation. In phase 1, we summarized the evaluation techniques used by the authors of the research papers, where 58% (30/52) evaluated performance. We used MARS evaluation in phase 2 and SonarCloud in phase 3. Thus, the results of the evaluation were different for each phase.

Table 2. Characteristics of phase 1 studies (N=52).

Characteristics	Phase 1 studies, n (%)
Number of surveyed studies or apps	
Android	45 (87)
iOS	6 (12)
Both	1 (2)
Roles of ML^a	
Recognition	19 (37)
Detection	6 (12)
Prediction	4 (8)
Recognition and monitoring	4 (8)
Recognition and recommendation	4 (8)
Recognition and estimation	2 (4)
Recommendation and monitoring	1 (2)
Recommendation	1 (2)
Estimation	1 (2)
Recognition, recommendation, and monitoring	1 (2)
Types of ML	
Supervised learning	39 (75)
Unsupervised learning	1 (2)
Both	1 (2)
External ML library	2 (4)
Processing techniques	
Data	25 (48)
Image	11 (21)
Image and calculation	4 (8)
Data and calculation	4 (8)
Voice	3 (6)
Calculation	3 (6)
Image, data, and calculation	2 (4)
Focus	
Physical health	19 (37)
Weight control	14 (27)
Disease	9 (17)
Mental health	5 (10)
Sleep	3 (6)
Recipe's recommendation	1 (2)
Multidimensional	1 (2)
Crucial functionalities	
Recognition	20 (38)
Detection	9 (17)
Prediction	4 (8)
Recognition and recommendation	4 (8)
Recognition and monitoring	4 (8)

Characteristics	Phase 1 studies, n (%)
Recommendation and monitoring	3 (6)
Monitoring	3 (6)
Recognition and estimation	2 (4)
Estimation	1 (2)
Recommendation	1 (2)
Recognition, recommendation, and monitoring	1 (2)
UI^b components	
Label	18 (35)
Image	17 (33)
Button	15 (29)
Input box	8 (15)
List	8 (15)
Navigation and structure	
Tab	6 (12)
Back and next	5 (10)
Main page	2 (4)
Tab and back and next	1 (2)
Tab, back and next, and hamburger menu	1 (2)
Services and technologies	
Motion sensors	21 (40)
Camera	18 (35)
GPS	2 (4)
Microphone	4 (8)
Security features	
Log-in	1 (2)
Architectures and patterns	
Client-server (web-based)	30 (58)
On device (offline)	19 (37)
MVC ^c	2 (4)
Inference of ML	
Web-based inference	18 (35)
Offline inference	10 (19)
Both	3 (7)
Web-ready solutions	2 (4)
Development approach	
Prototype	8 (15)
User-centered design	2 (4)
Agile	1 (2)
Extreme programming	1 (2)
Iterative	1 (2)
Logging mechanisms	
Automatic	24 (46)

Characteristics	Phase 1 studies, n (%)
Manual	24 (46)
Both	4 (8)
Evaluation	
Algorithm's performance	30 (58)
Algorithm's accuracy	9 (17)
Algorithm's performance and cross-validation	8 (15)
Usability study	4 (8)
Cross-validation	1 (2)

^aML: machine learning.

^bUI: user interface.

^cMVC: model-view-controller.

Table 3. Characteristics of phase 2 studies (N=42).

Characteristics	Phase 2 studies, n (%)
Number of surveyed studies or apps	
iOS	21 (50)
Android	21 (50)
Roles of ML^a	
Recognition, monitoring, and personalization	6 (14)
Monitoring and personalization	6 (14)
Recognition	2 (5)
Processing techniques	
Calculation	22 (52)
Calculation and data	8 (19)
Calculation, data, and image	6 (14)
Voice	2 (5)
Focus	
Physical health	16 (38)
Weight control	12 (29)
Women's health	6 (14)
Sleep	6 (14)
Behavior change	2 (5)
Crucial functionalities	
Monitoring	22 (52)
Recognition, monitoring, and personalization	6 (14)
Monitoring and personalization	6 (14)
Recognition	2 (5)
UI^b components	
Label	42 (100)
Image	42 (100)
Button	42 (100)
List	42 (100)
Scroll bar	42 (100)
Input box	34 (81)
Navigation and structure	
Tab (iOS)	18 (43)
Tab (Android)	17 (40)
Main page and hamburger menu (Android)	3 (7)
Tab and hamburger menu (iOS)	2 (5)
Main page and hamburger menu (iOS)	1 (2)
Main page (Android)	1 (2)
Services and technologies	
Camera	20 (48)
GPS	26 (62)
Motion sensors	7 (17)

Characteristics	Phase 2 studies, n (%)
Microphone	4 (10)
Security features	
Log-in	40 (95)
Logging mechanisms	
Manual	26 (62)
Automatic	10 (24)
Both	6 (14)
Evaluation	
MARS ^c	42 (100)

^aML: machine learning.

^bUI: user interface.

^cMARS: Mobile App Rating Scale.

Table 4. Characteristics of phase 3 studies (N=24).

Characteristics	Phase 3 studies, n (%)
Number of surveyed studies or apps	
iOS	13 (54)
Android	11 (46)
Roles of ML^a	
Recommendation and monitoring	2 (8)
Recognition	2 (8)
Recognition and recommendation	2 (8)
Recognition and monitoring	1 (4)
Types of ML	
External ML library	7 (29)
Processing techniques	
Calculation	13 (54)
Calculation and data	5 (21)
Data	2 (8)
Image	1 (4)
Voice	1 (4)
Calculation and image	1 (4)
Focus	
Weight control	7 (29)
Physical health	6 (25)
Monitoring	4 (17)
Mental health	3 (13)
Women's health	2 (8)
Behavior change	1 (4)
Multidimensional	1 (4)
Crucial functionalities	
Monitoring	13 (54)
Recommendation and monitoring	4 (17)
Recognition	2 (8)
Recognition and recommendation	2 (8)
Recognition and monitoring	1 (4)
Monitoring and personalization	1 (4)
UI^b components	
Label	24 (100)
Input box	23 (96)
Image	22 (92)
Button	22 (92)
List	16 (67)
Navigation and structure	
Tab (iOS)	11 (46)
Tab (Android)	4 (17)

Characteristics	Phase 3 studies, n (%)
Main page and menu (Android)	3 (13)
Main page (Android)	3 (13)
Main page (iOS)	1 (4)
Tab and hamburger menu (iOS)	1 (4)
Tab and hamburger menu (Android)	1 (4)
Services and technologies	
GPS	7 (29)
Camera	5 (21)
Motion sensors	1 (4)
Microphone	1 (4)
Security features	
Log-in	10 (42)
Architectures and patterns	
Client-server (web-based)	20 (83)
On device (offline)	4 (17)
MVC ^c	18 (75)
MVVM ^d	5 (21)
VIPER ^e	1 (4)
Inference of ML	
Web-based inference through ready solutions	7 (29)
Logging mechanisms	
Manual	18 (75)
Both	4 (17)
Automatic	2 (8)
Evaluation	
SonarCloud	24 (100)

^aML: machine learning.

^bUI: user interface.

^cMVC: model-view-controller.

^dMVVM: model-view-viewmodel.

^eVIPER: view-interactor-presenter-entity-router.

In terms of challenges and issues, phase 1 mentioned several challenges, such as the restricted number of resources in mobile devices, mobile device fragmentation, security of mobile apps, drawbacks of web-based and offline mechanisms, designing of an attractive and sustainable mobile app, and difficulties in collecting and processing data to design an efficient ML algorithm. Phase 2 reported the importance of network connectivity, which may affect the efficiency of a mobile app in the case of a connection loss. In addition, it highlighted the issues of writing separate implementations for each platform and the absence of care provider involvement in the development and evaluation phases. Of the analyzed apps in phase 3, we found that iOS apps had more bugs than Android apps, which may, however, be because of the developers and not the platform. Most apps on both platforms had code smells,

and performance issues. Only iOS apps had vulnerability issues. Furthermore, the logging mechanisms of the second and third phases were primitive and needed improvement to remain up to date with those described in research papers.

As a result, we found that commercial apps in phase 2 and open-source apps in phase 3 had more common aspects than the apps of research papers in phase 1. They were similar in that most used calculation methods as processing techniques and monitoring as a crucial functionality. In addition, they were simple and complete apps that were partially supported by ML and automatic logging. In contrast, apps in phase 1 were complex and intelligent, although some of them were incomplete, presenting a gap between real and research paper apps.

Principal Findings

This research involved various studies and apps designed for the general population with the aim of self-health care management. Most of these apps were developed with a specific focus, requiring users to download several apps to cover different aspects. Therefore, multidimensional well-being apps that combine multiple focuses need more research and development as it is better to download a single app with a set of features than to download several apps.

Furthermore, we found that the development of self-management mHealth apps required significant efforts from researchers to build and evaluate new algorithms and from developers to deal with different techniques and frequent updates of apps to stay up to date with the latest technologies. A parent example is developing apps with ML algorithms, which comprise several steps implemented manually, including collecting data, extracting features, and applying several ML algorithms to determine the most accurate algorithm. For example, the authors of S7 compared 7 classifiers: support vector machines, naïve Bayes, K-nearest neighbor, decision trees, LR, NNs, and rule-based classifiers. Furthermore, many other studies manually compared multiple algorithms to find the best algorithm, requiring a long time and great effort from researchers. Another example is the development of the same app in different languages and techniques to be compatible with multiple platforms.

In contrast, mobile devices are handheld gadgets with limited resources (eg, storage, computational power, and battery energy), which significantly hinders the improvement of service qualities such as ML algorithms that require dealing with intensive data and heavy computations. Connection with remote services such as the cloud can address these limitations. Therefore, many apps that used ML algorithms followed a web-based inference to achieve optimal performance within a reasonable time. However, this approach is generally insufficient when the connection is lost and may pose security issues. Therefore, some apps integrated a pretrained model with the mobile app (offline inference), which may cause some difficulty when updating the models. Only 4% (2/52) of apps supported both web-based and offline inferences for the same functionality but with lower performance and high energy consumption. Consequently, many challenges still exist related to finding an adequate algorithm that fulfills the specific requirements of intelligent self-management mHealth apps, as well as an efficient architecture that supports web-based and offline inferences with the ability to reduce resource consumption and execution time and increase performance, specifically when using large ML algorithms.

Threats to Validity

Threats to Internal Validity

Instrumental Bias

To ensure the consistency of our evaluation results, all evaluation processes in the second and third phases were performed in the same manner by the first author. The evaluation in phase 2 was applied to the same app on both Android and

iOS platforms. Furthermore, the evaluation process of phase 3 was repeated to double check the results.

Selection Bias

To ensure that we adopted unbiased and consistent procedures in the selection, we used the quasi-gold standard approach [4], which includes manual and automated search strategies, as well as snowballing. We selected the highest-quality peer-reviewed papers published in 7 web-based digital libraries. We further complemented our research with snowballing to capture as many studies as possible and minimize the potential for missing any relevant studies.

Threats to External Validity: Generalization to Different Samples

We reviewed studies that involved the implementation of a research study from 2008 to 2020. However, the generalizability of our findings could be affected by the exclusion of studies that presented theoretical research without implementation, as well as studies and apps linked with external devices.

Threats to Construct Validity

The RQs of this review could not entirely cover all of the reviewed research papers and self-management mHealth apps. Some research papers and apps had fewer or more details than the information identified in our RQs.

Threats to Content Validity

Relevance

To comprehensively identify the characteristics and issues of the selected studies, we divided the review into 3 main phases using different data sources, including research papers, commercial apps from digital Apple and Android app stores, and open-source apps from GitHub.

Representativeness

In phases 1 and 3, we selected mobile apps developed with either Kotlin or Swift on the Android and iOS platforms, respectively. These apps had a wide range of functionalities and purposes related to self-health care management.

Threats to Conclusion Validity

We extracted data from the assigned studies and self-management mHealth apps from the app stores and GitHub. To ensure the validity and consistency of the extracted data, the protocol for the data extraction strategy and format was developed by the first author and reviewed by the second author. In addition, we created a Microsoft Excel file to record and arrange the extracted data and check their relevance to our RQs.

Limitations and Future Work

This study had some limitations. The review was prepared and reviewed by 2 authors; however, it would have been better if it had more reviewers. In the second phase, apps were limited to self-health care management and included only those available in both Android and iOS digital app stores. The review would be more comprehensive if phase 2 included other categories of self-management mHealth apps, and phases 1 and 3 included other programming languages, such as Java and Objective-C. Furthermore, the review might have been more generalizable

if it included an app designed for children and people with special needs.

Developing mobile apps requires significant effort because of the complexity of self-management mHealth apps. Therefore, we have started the development of a framework that accelerates and facilitates the development of mHealth apps [82,83]. The framework semiautomatically generates Android and iOS mobile apps and will be enhanced with frequent characteristics that resulted from this review, such as tab structure, predefined components of ML algorithms, and local and external services. Moreover, the framework supports both web-based and offline inference, which appears to be a limitation of current self-management mHealth apps that usually support one of them, as the web-based mechanism could lead to unusable apps if the connection is lost, whereas the offline mechanism requires updating the entire app with each algorithm enhancement and library update.

Comparison With Prior Work

Several SLRs have been conducted on mHealth apps. For example, in the study by Mosa et al [84], the authors classified the functionalities of mHealth apps. They found that smartphones were useful tools for self-health care and clinical communication. Furthermore, smartphones can be used for the remote monitoring of patients, disease self-management, and patient education. In the study by Dounavi and Tsoumani [85], the authors described the effectiveness of mHealth apps in facilitating weight management behaviors by following healthy food consumption and physical activity. They found that mHealth apps are considered easy to use and useful in achieving weight loss because they involve users in the treatment plan, thereby increasing their commitment. These studies focused on mobile apps from a health care perspective. In contrast, our focus was on the software engineering perspective to identify the characteristics and challenges of mHealth apps, helping developers and researchers understand the infrastructure and functional and technical aspects of mHealth apps. Another SLR [86] focused on examining and identifying the empirical usability evaluation processes of mHealth apps. They stated that these processes could be improved by adopting automated mechanisms and combining >1 evaluation method. Furthermore, they demonstrated the importance of adapting mHealth apps according to user requirements.

This study aimed to conduct a comprehensive review and evaluation of self-management mHealth apps. This study gathered empirical evidence from the literature to identify the characteristics and challenges of existing self-management mHealth apps focused on self-health care management. The main contribution of this research is its detailed analysis and synthesis of relevant literature by software engineering researchers to deeply understand state of the art and provide guidance for the development of complex self-management mHealth apps.

Conclusions

In this research, we presented the details of an SLR on self-health care management mobile apps that consisted of three main phases. The results of this research can serve as a basis for researchers and developers to understand the characteristics of self-management mHealth apps and know the existing challenges that require further research. In phase 1, we reviewed 44 studies published between 2008 and 2020. In phase 2, 42 apps were reviewed and evaluated using the MARS. In phase 3, we reviewed and evaluated, using SonarCloud, 24 open-source apps from GitHub, including both iOS and Android platforms.

The research papers in phase 1 presented many interesting ideas, used different ML algorithms, and supported automatic logging mechanisms. These algorithms were used to process data to automatically recognize physical activities; diagnose diseases; recognize the types, quantities, and calories of food; and predict the user's emotion. However, the results of phase 1 show the need for optimization of the architecture and algorithm of intelligent self-management mHealth apps to efficiently include web-based and offline inferences, reduce resource consumption, and increase performance.

In phases 2 and 3, we found that most apps in app stores and GitHub focused on monitoring and analysis functionalities that use calculation methods to create progress reports and charts. However, the quantity of food consumed, exercise, and emotions should be entered manually. They lack automatic recognition of the type and quantity of food or activities. As a result, some advanced features exist in research papers but not in app stores and open-source apps, which may indicate that these features are still under research and development. Subsequently, the apps of phases 2 and 3 might need some improvement to keep pace with the advancement of research.

Acknowledgments

This work was part of a PhD dissertation funded by the University of Tabuk, Saudi Arabia.

Conflicts of Interest

None declared.

Multimedia Appendix 1

Context of mobile health apps.

[[DOCX File, 39 KB - humanfactors_v9i2e29767_app1.docx](#)]

Multimedia Appendix 2

Characteristics of mobile health apps.

[\[DOCX File, 27 KB - humanfactors_v9i2e29767_app2.docx\]](#)

Multimedia Appendix 3

Mobile App Rating Scale evaluation of phase 2.

[\[DOCX File, 31 KB - humanfactors_v9i2e29767_app3.docx\]](#)

Multimedia Appendix 4

SonarCloud evaluation of phase 3.

[\[DOCX File, 15 KB - humanfactors_v9i2e29767_app4.docx\]](#)**References**

1. Kitchenham B, Charters S. Guidelines for performing systematic literature reviews in software engineering. Keele University and Durham University Joint Report. 2007. URL: https://www.elsevier.com/_data/promis_misc/525444systematicreviewsguide.pdf [accessed 2021-04-04]
2. Stoyanov SR, Hides L, Kavanagh DJ, Zelenko O, Tjondronegoro D, Mani M. Mobile app rating scale: a new tool for assessing the quality of health mobile apps. *JMIR Mhealth Uhealth* 2015 Mar 11;3(1):e27 [FREE Full text] [doi: [10.2196/mhealth.3422](https://doi.org/10.2196/mhealth.3422)] [Medline: [25760773](https://pubmed.ncbi.nlm.nih.gov/25760773/)]
3. Sonarcloud. URL: <https://sonarcloud.io> [accessed 2021-04-04]
4. Zhang H, Babar MA, Tell P. Identifying relevant studies in software engineering. *Inf Softw Technol* 2011 Jun;53(6):625-637 [FREE Full text] [doi: [10.1016/j.infsof.2010.12.010](https://doi.org/10.1016/j.infsof.2010.12.010)]
5. iPhone app store downloads top 10 million in first weekend. Apple Newsroom. 2008 Jun 14. URL: <https://www.apple.com/newsroom/2008/07/14iPhone-App-Store-Downloads-Top-10-Million-in-First-Weekend/> [accessed 2021-04-04]
6. Li H, Zhang Q, Lu K. Integrating mobile sensing and social network for personalized health-care application. In: Proceedings of the 30th Annual ACM Symposium on Applied Computing. 2015 Presented at: SAC '15; April 13-17, 2015; Salamanca, Spain p. 527-534 URL: <https://doi.org/10.1145/2695664.2695767> [doi: [10.1145/2695664.2695767](https://doi.org/10.1145/2695664.2695767)]
7. Jia B, Li J. Recognizing human activities in real-time using mobile phone sensors. In: Proceedings of the 8th China Conference on Advances in Wireless Sensor Networks. 2014 Presented at: CWSN '14; October 31-November 2, 2014; Xi'an, China p. 638-650 URL: https://link.springer.com/chapter/10.1007/978-3-662-46981-1_60 [doi: [10.1007/978-3-662-46981-1_60](https://doi.org/10.1007/978-3-662-46981-1_60)]
8. Guo J, Zhou X, Sun Y, Ping G, Zhao G, Li Z. Smartphone-based patients' activity recognition by using a self-learning scheme for medical monitoring. *J Med Syst* 2016 Jun;40(6):140. [doi: [10.1007/s10916-016-0497-2](https://doi.org/10.1007/s10916-016-0497-2)] [Medline: [27106584](https://pubmed.ncbi.nlm.nih.gov/27106584/)]
9. Wang C, Xu Y, Zhang J, Yu W. COPO: a novel position-adaptive method for smartphone-based human activity recognition. In: Proceedings of the 10th Asia-Pacific Services Computing Conference on Advances in Services Computing. 2016 Presented at: APSCC '16; November 16-18, 2016; Zhangjiajie, China p. 1-14 URL: https://link.springer.com/chapter/10.1007/978-3-319-49178-3_1 [doi: [10.1007/978-3-319-49178-3_1](https://doi.org/10.1007/978-3-319-49178-3_1)]
10. Prudêncio J, Aguiar A, Lucani D. Physical activity recognition from smartphone Embedded Sensors. In: Proceedings of the 6th Iberian Conference on Pattern Recognition and Image Analysis. 2013 Presented at: IbPRIA '13; June 5-7, 2013; Madeira, Portugal p. 863-872 URL: https://link.springer.com/chapter/10.1007/978-3-642-38628-2_102 [doi: [10.1007/978-3-642-38628-2_102](https://doi.org/10.1007/978-3-642-38628-2_102)]
11. Lu Y, Wei Y, Liu L, Zhong J, Sun L, Liu Y. Towards unsupervised physical activity recognition using smartphone accelerometers. *Multimed Tools Appl* 2016 Jan 9;76(8):10701-10719 [FREE Full text] [doi: [10.1007/s11042-015-3188-y](https://doi.org/10.1007/s11042-015-3188-y)]
12. Shoaib M, Scholten H, Havinga PJ. Towards physical activity recognition using smartphone sensors. In: Proceedings of the IEEE 10th International Conference on Ubiquitous Intelligence and Computing and IEEE 10th International Conference on Autonomic and Trusted Computing. 2013 Presented at: UIC-ATC '13; December 18-21, 2013; Vietri sul Mare, Italy p. 80-87. [doi: [10.1109/uic-atc.2013.43](https://doi.org/10.1109/uic-atc.2013.43)]
13. Cao L, Wang Y, Jin Q, Ma J. ActiRecognizer: design and implementation of a real-time human activity recognition system. In: Proceedings of the 2017 International Conference on Cyber-Enabled Distributed Computing and Knowledge Discovery. 2017 Presented at: CyberC '17; October 12-14, 2017; Nanjing, China p. 266-271. [doi: [10.1109/CyberC.2017.71](https://doi.org/10.1109/CyberC.2017.71)]
14. Chen Z, Jiang C, Xie L. A novel ensemble ELM for human activity recognition using smartphone sensors. *IEEE Trans Ind Inf* 2019 May;15(5):2691-2699. [doi: [10.1109/tii.2018.2869843](https://doi.org/10.1109/tii.2018.2869843)]
15. Beily MD, Badjowawo MD, Bekak DO, Dana S. A sensor based on recognition activities using smartphone. In: Proceedings of the 2016 International Seminar on Intelligent Technology and Its Applications. 2016 Presented at: ISITIA '16; July 28-30, 2016; Lombok, Indonesia p. 393-398. [doi: [10.1109/isitia.2016.7828692](https://doi.org/10.1109/isitia.2016.7828692)]
16. Saha J, Chowdhury C, Biswas S. Device independent activity monitoring using smart handhelds. In: Proceedings of the 7th International Conference on Cloud Computing, Data Science & Engineering. 2017 Presented at: Confluence '17; January 12-13, 2017; Noida, India p. 406-411. [doi: [10.1109/CONFLUENCE.2017.7943184](https://doi.org/10.1109/CONFLUENCE.2017.7943184)]

17. Ravi D, Lo B, Yang GZ. Real-time food intake classification and energy expenditure estimation on a mobile device. In: Proceedings of the IEEE 12th International Conference on Wearable and Implantable Body Sensor Networks. 2015 Presented at: BSN '15; June 9-12, 2015; Cambridge, MA, USA p. 1-6. [doi: [10.1109/bsn.2015.7299410](https://doi.org/10.1109/bsn.2015.7299410)]
18. Peddi SV, Yassine A, Shirmohammadi S. Cloud based virtualization for a calorie measurement e-health mobile application. In: Proceedings of the 2015 IEEE International Conference on Multimedia & Expo Workshops. 2015 Presented at: ICMEW '15; June 29-July 3, 2015; Turin, Italy p. 1-6. [doi: [10.1109/icmew.2015.7169853](https://doi.org/10.1109/icmew.2015.7169853)]
19. Villalobos G, Almaghrabi R, Pouladzadeh P, Shirmohammadi S. An image processing approach for calorie intake measurement. In: Proceedings of the 2012 IEEE International Symposium on Medical Measurements and Applications. 2012 Presented at: MeMeA '12; May 18-19, 2012; Budapest, Hungary p. 1-5. [doi: [10.1109/MeMeA.2012.6226636](https://doi.org/10.1109/MeMeA.2012.6226636)]
20. Ocaý AB, Fernandez JM, Palaoag TD. NutriTrack: Android-based food recognition app for nutrition awareness. In: Proceedings of the 3rd IEEE International Conference on Computer and Communications. 2017 Presented at: ICCS '17; December 13-16, 2017; Chengdu, China p. 2099-2104. [doi: [10.1109/CompComm.2017.8322907](https://doi.org/10.1109/CompComm.2017.8322907)]
21. Tangsrapiroj S, Wongkham N, Leelalerkiat B, Chuenpukdi S. WhatTheHealth: an Android application for consumers of healthy food. In: Proceedings of the 16th International Joint Conference on Computer Science and Software Engineering. 2019 Presented at: JCSSE '19; July 10-12, 2019; Chonburi, Thailand p. 61-66. [doi: [10.1109/JCSSE.2019.8864192](https://doi.org/10.1109/JCSSE.2019.8864192)]
22. Kumbhar D, Patil S. Mobile cloud based system recognizing nutrition and freshness of food image. In: Proceedings on the 2017 International Conference on Energy, Communication, Data Analytics and Soft Computing. 2017 Presented at: ICECDS '17; August 1-2, 2017; Chennai, India p. 709-714. [doi: [10.1109/ICECDS.2017.8389528](https://doi.org/10.1109/ICECDS.2017.8389528)]
23. Montanini L, Sabino N, Spinsante S, Gambi E. Smartphone as unobtrusive sensor for real-time sleep recognition. In: Proceedings of the 2018 IEEE International Conference on Consumer Electronics. 2018 Presented at: ICCE '18; January 12-14, 2018; Las Vegas, NV, USA p. 1-4. [doi: [10.1109/ICCE.2018.8326220](https://doi.org/10.1109/ICCE.2018.8326220)]
24. Yamashita Y, Onodera M, Shimoda K, Tobe Y. Emotion-Polarity Visualizer on Smartphone. In: Proceedings of the IEEE 27th International Requirements Engineering Conference Workshops. 2019 Presented at: REW '19; September 23-27, 2019; Jeju, South Korea p. 96-99. [doi: [10.1109/REW.2019.00020](https://doi.org/10.1109/REW.2019.00020)]
25. Shrestha A, Won M. DeepWalking: enabling smartphone-based walking speed estimation using deep learning. In: Proceedings of the 2018 IEEE Global Communications Conference. 2018 Presented at: GLOBECOM '18; December 9-13, 2018; Abu Dhabi, United Arab Emirates p. 1-6. [doi: [10.1109/GLOCOM.2018.8647857](https://doi.org/10.1109/GLOCOM.2018.8647857)]
26. Matsumura K, Yamakoshi T. iPhysioMeter: a new approach for measuring heart rate and normalized pulse volume using only a smartphone. *Behav Res Methods* 2013 Dec;45(4):1272-1278. [doi: [10.3758/s13428-012-0312-z](https://doi.org/10.3758/s13428-012-0312-z)] [Medline: [23408381](https://pubmed.ncbi.nlm.nih.gov/23408381/)]
27. Pelegris P, Banitsas K, Orbach T, Marias K. A novel method to detect heart beat rate using a mobile phone. *Annu Int Conf IEEE Eng Med Biol Soc* 2010;2010:5488-5491. [doi: [10.1109/IEMBS.2010.5626580](https://doi.org/10.1109/IEMBS.2010.5626580)] [Medline: [21096290](https://pubmed.ncbi.nlm.nih.gov/21096290/)]
28. Dai X, Spasić I, Meyer B, Chapman S, Andres F. Machine learning on mobile: an on-device inference app for skin cancer detection. In: Proceedings of the 4th International Conference on Fog and Mobile Edge Computing. 2019 Presented at: FMEC '19; June 10-13, 2019; Rome, Italy p. 301-305. [doi: [10.1109/FMEC.2019.8795362](https://doi.org/10.1109/FMEC.2019.8795362)]
29. Ech-Cherif A, Misbhauddin M, Ech-Cherif M. Deep neural network based mobile dermoscopy application for triaging skin cancer detection. In: Proceedings of the 2nd International Conference on Computer Applications & Information Security. 2019 Presented at: ICCAIS '19; May 1-3, 2019; Riyadh, Saudi Arabia p. 1-6. [doi: [10.1109/CAIS.2019.8769517](https://doi.org/10.1109/CAIS.2019.8769517)]
30. Wadhawan T, Situ N, Lancaster K, Yuan X, Zouridakis G. SkinScan©: a portable library for melanoma detection on handheld devices. *Proc IEEE Int Symp Biomed Imaging* 2011 Mar 30;2011:133-136 [FREE Full text] [doi: [10.1109/ISBI.2011.5872372](https://doi.org/10.1109/ISBI.2011.5872372)] [Medline: [21892382](https://pubmed.ncbi.nlm.nih.gov/21892382/)]
31. Ghandeharioun A, McDuff D, Czerwinski M, Rowan K. EMMA: an emotion-aware wellbeing chatbot. In: Proceedings of the 8th International Conference on Affective Computing and Intelligent Interaction. 2019 Presented at: ACII '19; September 3-6, 2019; Cambridge, UK p. 1-7. [doi: [10.1109/ACII.2019.8925455](https://doi.org/10.1109/ACII.2019.8925455)]
32. Jiménez-Serrano S, Tortajada S, García-Gómez JM. A mobile health application to predict postpartum depression based on machine learning. *Telemed J E Health* 2015 Jul;21(7):567-574. [doi: [10.1089/tmj.2014.0113](https://doi.org/10.1089/tmj.2014.0113)] [Medline: [25734829](https://pubmed.ncbi.nlm.nih.gov/25734829/)]
33. Figueiredo IN, Leal C, Pinto L, Bolito J, Lemos A. Exploring smartphone sensors for fall detection. *mUX J Mob User Exp* 2016 May 5;5(1):2 [FREE Full text] [doi: [10.1186/s13678-016-0004-1](https://doi.org/10.1186/s13678-016-0004-1)]
34. Bhatia A, Kumar S, Mago V. Gradient: a user-centric lightweight smartphone based standalone fall detection system. In: Proceedings of the 17th Portuguese Conference on Artificial Intelligence. 2015 Presented at: EPIA '15; September 8-11, 2015; Coimbra, Portugal p. 67-78 URL: https://doi.org/10.1007/978-3-319-23485-4_7 [doi: [10.1007/978-3-319-23485-4_7](https://doi.org/10.1007/978-3-319-23485-4_7)]
35. Sanabria P, Benedetto JI, Neyem A, Navon J, Poellabauer C. Code offloading solutions for audio processing in mobile healthcare applications: a case study. In: Proceedings of the IEEE/ACM 5th International Conference on Mobile Software Engineering and Systems. 2018 Presented at: MOBILESoft '18; May 27-June 3, 2018; Gothenburg, Sweden p. 117-121. [doi: [10.1145/3197231.3197256](https://doi.org/10.1145/3197231.3197256)]
36. Khan NS, Muaz MH, Kabir A, Islam MN. Diabetes predicting mHealth application using machine learning. In: Proceedings of the 2017 IEEE International WIE Conference on Electrical and Computer Engineering. 2017 Presented at: WIECON-ECE '17; December 18-19, 2017; Dehradun, India p. 237-240. [doi: [10.1109/WIECON-ECE.2017.8468885](https://doi.org/10.1109/WIECON-ECE.2017.8468885)]
37. Kelly D, Curran K, Caulfield B. Automatic prediction of health status using smartphone-derived behavior profiles. *IEEE J Biomed Health Inform* 2017 Nov;21(6):1750-1760. [doi: [10.1109/JBHI.2017.2649602](https://doi.org/10.1109/JBHI.2017.2649602)] [Medline: [28092582](https://pubmed.ncbi.nlm.nih.gov/28092582/)]

38. Fahim M, Vui LB, Fatima I, Lee S, Yoon Y. A sleep monitoring application for u-lifecare using accelerometer sensor of smartphone. In: Proceedings of the 7th International Conference on Ubiquitous Computing and Ambient Intelligence: Context-Awareness and Context-Driven Interaction. 2013 Presented at: UCAmI '13; December 2-6, 2013; Carrillo, Costa Rica p. 151-158 URL: https://doi.org/10.1007/978-3-319-03176-7_20 [doi: [10.1007/978-3-319-03176-7_20](https://doi.org/10.1007/978-3-319-03176-7_20)]
39. Sandulescu V, Andrews S, Ellis D, Dobrescu R, Martinez-Mozos O. Mobile app for stress monitoring using voice features. In: Proceedings of the 2015 E-Health and Bioengineering Conference. 2015 Presented at: EHB '15; November 19-21, 2015; Iasi, Romania p. 1-4. [doi: [10.1109/EHB.2015.7391411](https://doi.org/10.1109/EHB.2015.7391411)]
40. Kalaiselvan V, Azman F, Cheng LK, Rahim FA. nocturnOWL: sleep-monitoring virtual pet mobile application. In: Proceedings of the 2019 IEEE Conference on Open Systems. 2019 Presented at: ICOS '19; November 19-21, 2019; Pulau Pinang, Malaysia p. 51-56. [doi: [10.1109/ICOS47562.2019.8975709](https://doi.org/10.1109/ICOS47562.2019.8975709)]
41. Lomaliza JP, Park H. A highly efficient and reliable heart rate monitoring system using smartphone cameras. *Multimed Tools Appl* 2017;76(20):21051-21071 [FREE Full text] [doi: [10.1007/s11042-016-4010-1](https://doi.org/10.1007/s11042-016-4010-1)]
42. Yang PC, Su SC, Wu IL, Chiang JH. Life Record: a smartphone-based daily activity monitoring system. In: Proceedings of the 6th International Conference on Advances in Swarm and Computational Intelligence. 2015 Presented at: CCI '15; June 25-28, 2015; Beijing, China p. 380-385 URL: https://doi.org/10.1007/978-3-319-20472-7_41 [doi: [10.1007/978-3-319-20472-7_41](https://doi.org/10.1007/978-3-319-20472-7_41)]
43. Villarreal V, Otero M. Development a mobile system based on the harris-benedict equation to indicate the caloric intake. In: Proceedings of the 10th International Conference on Ubiquitous Computing and Ambient Intelligence. 2016 Presented at: UCAmI '16; November 29-December 2 2016; Gran Canaria, Spain p. 264-254 URL: https://doi.org/10.1007/978-3-319-48746-5_25 [doi: [10.1007/978-3-319-48746-5_25](https://doi.org/10.1007/978-3-319-48746-5_25)]
44. de la Torre Díez I, Garcia-Zapirain B, López-Coronado M, Rodrigues JJ, Del Pozo Vegas C. A new mHealth app for monitoring and awareness of healthy eating: development and user evaluation by Spanish users. *J Med Syst* 2017 Jul;41(7):109. [doi: [10.1007/s10916-017-0753-0](https://doi.org/10.1007/s10916-017-0753-0)] [Medline: [28555352](https://pubmed.ncbi.nlm.nih.gov/28555352/)]
45. Sánchez CH, Ramírez MR, Palencia JS, Lobato BY, del Carmen Osuna Millán N. My health info: an informative mHealth app for healthcare and weight control. In: Proceedings of the 5th KES International Conference on Innovation in Medicine and Healthcare. 2017 Presented at: KES-InMed '17; June 21-23, 2017; Vilamoura, Portugal p. 3-11 URL: https://doi.org/10.1007/978-3-319-59397-5_1 [doi: [10.1007/978-3-319-59397-5_1](https://doi.org/10.1007/978-3-319-59397-5_1)]
46. Harries T, Eslambolchilar P, Stride C, Rettie R, Walton S. Walking in the wild - using an always-on smartphone application to increase physical activity. In: Proceedings of the 14th IFIP TC 13 International Conference on Human-Computer Interaction. 2013 Presented at: INTERACT '13; September 2-6, 2013; Cape Town, South Africa p. 19-36 URL: https://doi.org/10.1007/978-3-642-40498-6_2 [doi: [10.1007/978-3-642-40498-6_2](https://doi.org/10.1007/978-3-642-40498-6_2)]
47. Rahman F, Henninger P, Kegley D, Sullivan K, Yoo J. Healthy hankerings: motivating adolescents to combat obesity with a mobile application. In: Proceedings of the 20th International Conference on Human-Computer Interaction. Interaction in Context. 2018 Presented at: HCI '18; July 15-20, 2018; Las Vegas, NV, USA p. 185-194 URL: https://doi.org/10.1007/978-3-319-91244-8_15 [doi: [10.1007/978-3-319-91244-8_15](https://doi.org/10.1007/978-3-319-91244-8_15)]
48. Zhao Z, Sun Z, Huang L, Guo H, Wang J, Xu H. iRun: a smartphone-based system to alert runners to warm up before running. In: Proceedings of the 11th International Conference on Wireless Algorithms, Systems, and Applications. 2016 Presented at: WASA '16; August 8-10, 2016; Bozeman, MT, USA p. 186-196 URL: https://doi.org/10.1007/978-3-319-42836-9_17 [doi: [10.1007/978-3-319-42836-9_17](https://doi.org/10.1007/978-3-319-42836-9_17)]
49. Fahim M, Baker T, Khattak AM, Alfandi O. Alert me: enhancing active lifestyle via observing sedentary behavior using mobile sensing systems. In: Proceedings of the IEEE 19th International Conference on e-Health Networking, Applications and Services. 2017 Presented at: HealthCom '17; October 12-15, 2017; Dalian, China p. 1-4. [doi: [10.1109/HealthCom.2017.8210838](https://doi.org/10.1109/HealthCom.2017.8210838)]
50. Kong F, He H, Raynor HA, Tan J. DietCam: multi-view regular shape food recognition with a camera phone. *Pervasive Mob Comput* 2015 May;19:108-121 [FREE Full text] [doi: [10.1016/j.pmcj.2014.05.012](https://doi.org/10.1016/j.pmcj.2014.05.012)]
51. Yang S, Zhou P, Duan K, Hossain MS, Alhamid MF. emHealth: towards emotion health through depression prediction and intelligent health recommender system. *Mobile Netw Appl* 2018;23(2):216-226 [FREE Full text] [doi: [10.1007/s11036-017-0929-3](https://doi.org/10.1007/s11036-017-0929-3)]
52. Zhang W, Yu Q, Siddiquie B, Divakaran A, Sawhney H. "Snap-n-Eat": food recognition and nutrition estimation on a smartphone. *J Diabetes Sci Technol* 2015 May;9(3):525-533 [FREE Full text] [doi: [10.1177/1932296815582222](https://doi.org/10.1177/1932296815582222)] [Medline: [25901024](https://pubmed.ncbi.nlm.nih.gov/25901024/)]
53. Maruyama T, Kawano Y, Yanai K. Real-time mobile recipe recommendation system using food ingredient recognition. In: Proceedings of the 2nd ACM International Workshop on Interactive Multimedia on Mobile and Portable Devices. 2012 Presented at: IMMPD '12; November 2, 2012; Nara, Japan p. 27-34 URL: <https://doi.org/10.1145/2390821.2390830> [doi: [10.1145/2390821.2390830](https://doi.org/10.1145/2390821.2390830)]
54. Duan P, Wang W, Zhang W, Gong F, Zhang P, Rao Y. Food image recognition using pervasive cloud computing. In: Proceedings of the 2013 IEEE International Conference on Green Computing and Communications and IEEE Internet of Things and IEEE Cyber, Physical and Social Computing. 2013 Presented at: GreenCom-iThings-CPSCOM '13; August 20-23, 2013; Beijing, China p. 1631-1637. [doi: [10.1109/GreenCom-iThings-CPSCOM.2013.296](https://doi.org/10.1109/GreenCom-iThings-CPSCOM.2013.296)]

55. Tran DN, Phan DD. Human activities recognition in Android smartphone using support vector machine. In: Proceedings of the 7th International Conference on Intelligent Systems, Modelling and Simulation. 2016 Presented at: ISMS '16; January 25-27, 2016; Bangkok, Thailand p. 64-68. [doi: [10.1109/ISMS.2016.51](https://doi.org/10.1109/ISMS.2016.51)]
56. Zhu F, Bosch M, Woo I, Kim S, Boushey CJ, Ebert DS, et al. The use of mobile devices in aiding dietary assessment and evaluation. *IEEE J Sel Top Signal Process* 2010 Aug;4(4):756-766 [FREE Full text] [doi: [10.1109/JSTSP.2010.2051471](https://doi.org/10.1109/JSTSP.2010.2051471)] [Medline: [20862266](https://pubmed.ncbi.nlm.nih.gov/20862266/)]
57. Lane ND, Lin M, Mohammad M, Yang X, Lu H, Cardone G, et al. BeWell: sensing sleep, physical activities and social interactions to promote wellbeing. *Mobile Netw Appl* 2014 Jan 9;19(3):345-359 [FREE Full text] [doi: [10.1007/s11036-013-0484-5](https://doi.org/10.1007/s11036-013-0484-5)]
58. Calcheck Health App. GitHub. 2020 Jun 6. URL: <https://github.com/amaygada/Calcheck-HealthApp> [accessed 2021-04-14]
59. Corona Tracker Awareness App. GitHub. 2020 Apr 16. URL: https://github.com/sunmughan/corona_tracker-awareness_app [accessed 2021-04-14]
60. Easy Healthy Diet. GitHub. 2020 Sep 21. URL: https://github.com/Miltentt/Easy_Healthy_Diet [accessed 2021-04-14]
61. Eat Healthy App. GitHub. 2019 Nov 10. URL: <https://github.com/IS3261-team-awesome/EatHealthyApp> [accessed 2021-04-14]
62. Health Care. GitHub. 2019 Feb 8. URL: <https://github.com/jarupulapremkumar/healthcare-> [accessed 2021-04-14]
63. Health Diary. GitHub. 2020 Apr 26. URL: <https://github.com/gbarreiro/healthdiary> [accessed 2021-04-14]
64. jHealth. GitHub. 2020 May 25. URL: <https://github.com/bmarko98/jHealth> [accessed 2021-04-14]
65. LindaJamii. GitHub. 2020 Jun 29. URL: <https://github.com/polojerry/Linda-Jamii> [accessed 2021-04-14]
66. Period. GitHub. 2020 Jul 1. URL: <https://github.com/Sekushy/Period> [accessed 2021-04-14]
67. QuitSmoking. GitHub. 2020 Jul 5. URL: <https://github.com/scoute-dich/QuitSmoking> [accessed 2021-04-14]
68. SCD-Food-List-Android. GitHub. 2021 May 16. URL: <https://github.com/GeorgeMcLellan/SCD-Food-List-Android> [accessed 2021-04-14]
69. ActiveBoomers. GitHub. 2021 Jan 8. URL: <https://github.com/kartbnb/ActiveBoomers> [accessed 2021-04-14]
70. BodyMindfulness. GitHub. 2019 May 27. URL: <https://github.com/MartinezBlaze/BodyMindfulness> [accessed 2021-04-14]
71. CalorieCounter. GitHub. 2020 Apr 21. URL: <https://github.com/nsaigal/CalorieCounter> [accessed 2021-04-14]
72. Fitness. GitHub. 2020 May 10. URL: <https://github.com/nickhuang25/Fitness> [accessed 2021-04-14]
73. Health4Food. GitHub. 2019 Apr 15. URL: <https://github.com/wonathanjong/Health4Food> [accessed 2021-04-14]
74. HealthAppYouFit-Ios. GitHub. 2018 May 1. URL: <https://github.com/Tapadyuti1991/HealthAppYouFit-Ios> [accessed 2021-04-14]
75. HealthMate. GitHub. 2020 Jun 23. URL: <https://github.com/ohmpatel1997/HealthMate> [accessed 2021-04-14]
76. IOS-HealthTracker. GitHub. 2020 Jun 10. URL: <https://github.com/cakarakif/IOS-HealthTracker> [accessed 2021-04-14]
77. Healthy-Diet. GitHub. 2019 May 16. URL: <https://github.com/tertl700/Healthy-Diet> [accessed 2021-04-14]
78. iHealth. GitHub. 2019 Feb 13. URL: <https://github.com/rcasanovan/iHealth> [accessed 2021-04-14]
79. mHealthApp. GitHub. 2018 Sep 8. URL: <https://github.com/baksha97/mHealth> [accessed 2021-04-14]
80. Moodify. GitHub. 2019 Jan 17. URL: <https://github.com/richardxie68/Moodify> [accessed 2021-04-14]
81. smARt_pet. GitHub. 2019 Aug 15. URL: https://github.com/wwkb123/smARt_pet [accessed 2021-04-14]
82. Alwakeel L, Lano K. Model driven development of mobile application. King's College London. 2020 Sep 1. URL: https://kclpure.kcl.ac.uk/portal/files/141977701/Model_Driven_Development_of_Mobile_Applications_ECOOP20.pdf [accessed 2021-04-14]
83. Lano K, Alwakeel L, Rahimi SK, Haughton H. Synthesis of mobile applications using AgileUML. In: Proceedings of the 14th Innovations in Software Engineering Conference (formerly known as India Software Engineering Conference). 2021 Presented at: ISEC '21; February 25-27, 2021; Bhubaneswar, India p. 1-10 URL: <https://doi.org/10.1145/3452383.3452409> [doi: [10.1145/3452383.3452409](https://doi.org/10.1145/3452383.3452409)]
84. Mosa AS, Yoo I, Sheets L. A systematic review of healthcare applications for smartphones. *BMC Med Inform Decis Mak* 2012 Jul 10;12:67 [FREE Full text] [doi: [10.1186/1472-6947-12-67](https://doi.org/10.1186/1472-6947-12-67)] [Medline: [22781312](https://pubmed.ncbi.nlm.nih.gov/22781312/)]
85. Dounavi K, Tsoumani O. Mobile health applications in weight management: a systematic literature review. *Am J Prev Med* 2019 Jun;56(6):894-903 [FREE Full text] [doi: [10.1016/j.amepre.2018.12.005](https://doi.org/10.1016/j.amepre.2018.12.005)] [Medline: [31003801](https://pubmed.ncbi.nlm.nih.gov/31003801/)]
86. Zapata BC, Fernández-Alemán JL, Idri A, Toval A. Empirical studies on usability of mHealth apps: a systematic literature review. *J Med Syst* 2015 Feb;39(2):1. [doi: [10.1007/s10916-014-0182-2](https://doi.org/10.1007/s10916-014-0182-2)] [Medline: [25600193](https://pubmed.ncbi.nlm.nih.gov/25600193/)]

Abbreviations

- API:** application programming interface
LR: logistic regression
MARS: Mobile App Rating Scale
mHealth: mobile health
ML: machine learning
MVC: model-view-controller

NN: neural network
OS: operating system
RQ: research question
SLR: systematic literature review
UI: user interface

Edited by A Kushniruk; submitted 19.04.21; peer-reviewed by M Walrave, V Strotbaum; comments to author 22.06.21; revised version received 11.07.21; accepted 19.04.22; published 25.05.22.

Please cite as:

Alwakeel L, Lano K

Functional and Technical Aspects of Self-management mHealth Apps: Systematic App Search and Literature Review

JMIR Hum Factors 2022;9(2):e29767

URL: <https://humanfactors.jmir.org/2022/2/e29767>

doi: [10.2196/29767](https://doi.org/10.2196/29767)

PMID: [35612887](https://pubmed.ncbi.nlm.nih.gov/35612887/)

©Lyan Alwakeel, Kevin Lano. Originally published in JMIR Human Factors (<https://humanfactors.jmir.org>), 25.05.2022. This is an open-access article distributed under the terms of the Creative Commons Attribution License (<https://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Human Factors, is properly cited. The complete bibliographic information, a link to the original publication on <https://humanfactors.jmir.org>, as well as this copyright and license information must be included.

Review

Assessing the Legal Aspects of Information Security Requirements for Health Care in 3 Countries: Scoping Review and Framework Development

Prosper Kandabongee Yeng^{1*}, MSc; Muhammad Ali Fauzi^{1*}, MSc; Luyi Sun^{1*}, MSc; Bian Yang^{1*}, PhD

Department of Information Security and Communication Technology, Norwegian University of Science and Technology, Gjøvik, Norway

* all authors contributed equally

Corresponding Author:

Prosper Kandabongee Yeng, MSc
Department of Information Security and Communication Technology
Norwegian University of Science and Technology
Teknologivegen 22
Gjøvik, 2815
Norway
Phone: 47 96992743
Email: prosper.yeng@ntnu.no

Abstract

Background: The loss of human lives from cyberattacks in health care is no longer a probabilistic quantification but a reality that has begun. In addition, the threat scope is also expanding to involve a threat of national security, among others, resulting in surging data breaches within the health care sector. For that matter, there have been provisions of various legislation, regulations, and information security governance tools such as policies, standards, and directives toward enhancing health care information security-conscious care behavior among users. Meanwhile, in a research scenario, there are no comprehensive required security practices to serve as a yardstick in assessing security practices in health care. Moreover, an analysis of the holistic view of the requirements that need more concentration of management, end users, or both has not been comprehensively developed. Thus, there is a possibility that security practice research will leave out vital requirements.

Objective: The objective of this study was to systematically identify, assess, and analyze the state-of-the-art information security requirements in health care. These requirements can be used to develop a framework to serve as a yardstick for measuring the future real security practices of health care staff.

Methods: A scoping review was, as a result, adopted to identify, assess, and analyze the information security requirement sources within health care in Norway, Indonesia, and Ghana.

Results: Of 188 security and privacy requirement sources that were initially identified, 130 (69.1%) were fully read by the authors. Subsequently, of these 188 requirement documents, 82 (43.6%) fully met the inclusion criteria and were accessed and analyzed. In total, 253 security and privacy requirements were identified in this work. The findings were then used to develop a framework to serve as a benchmark for modeling and analyzing health care security practices.

Conclusions: On the basis of these findings, a framework for modeling, analyzing, and developing effective security countermeasures, including incentivization measures, was developed. Following this framework, research results of health care security practices would be more reliable and effective than relying on incomprehensive security requirements.

(*JMIR Hum Factors* 2022;9(2):e30050) doi:[10.2196/30050](https://doi.org/10.2196/30050)

KEYWORDS

legal requirement; information security; healthcare; security practice

Introduction

Background

There have been enormous gains in the application of information technology (IT) in health care in various areas such as decision support, telemedicine, electronic health record (EHR) management, chronic disease management with medical devices, drugs, and vaccine production [1-3]. However, cyberattacks in health care and their related adverse impact are a significant problem, especially in the midst of the infamous COVID-19 pandemic [4]. For example, Brno University Hospital in the Czech Republic was recently attacked, and cyberattackers were believed to have used spear phishing to gain access and deployed ransomware, which encrypted the data in the entire hospital network [5]. The hospital was compelled to shut down and battle with the cyberattack to restore its data. Even though the hospital was one of the COVID-19 treatment centers, the incident apparently prevented them from providing health care services during the attack period. Following that, there were other cyberattacks on the World Health Organization, Hammersmith Medicines Research Group in the United Kingdom (a COVID-19 vaccine trial group), the US Health and Human Services Department, Paris Hospital Authority in France, Bam Construct and Interserve (a COVID-19 hospital construction company), and Babylon Health (a hospital appointment and teleconsultation videoconferencing system) in the United Kingdom [6].

In addition, cybersecurity and privacy issues in health care have become a global concern as data breaches in health care continue to surge. In 2017, approximately 5 million health care records were compromised globally [1-3]. This tripled in 2018 to approximately 15 million, and the number of compromised health care records continues to increase yearly [3]. In addition, the cost associated with data breaches (eg, cost of detection of breaches, cost of fines paid in data breaches, cost of recovery, and payment of ransoms) is the highest in health care among various industries [7].

Data breaches and security issues in health care have major consequences on confidentiality, integrity, and availability (CIA). This usually perturbs the data subjects, the health care organizations, and the laws of the countries involved [8,9]. The adverse impact on data subjects includes situations in which the stolen data can be used as a means of pressure to demand other goals by criminals. Recently, an instance occurred in Finland [10], where stolen medical records were used by cybercriminals to pressure the data owners for money. Unauthorized persons can also disrupt the proper functioning of health care operations, such that the net effect can result in the loss of a patient's life. A related instance occurred in Germany, where a hospital's IT systems were hit by ransomware, which resulted in the death of a patient due to the unavailability of the health care system at the time of need [11]. Mutual trust and confidentiality between health care providers and patients [12-15], economic losses [10,15,16], privacy issues [9,17], and unreliable medical records [11,18,19] or medical devices [3] are some of the effects often faced by data subjects during cyberattacks in health care. It could be much

disheartening for patients to battle against their medical conditions, and at the same time, they have to battle with their privacy issues arriving from cyberattacks. Mutual trust with data between health care professionals and patients is very cardinal in terms of good-quality health provision. Health care professionals depend on the accuracy and comprehensiveness of the information provided by patients for therapeutic measures [13]. Therefore, health care providers are required to store large quantities of sensitive personal information of patients [14]. Similarly, patients trust that their personal information disclosed for medical reasons is to be kept confidential [15]. Sadly, this mutual trust in relation to patients' data is often broken in data breaches [15,16].

Furthermore, health care systems are targeted for various computer crimes with the intention of stealing, altering, hindering, and disrupting data or other functions [5,11]. The consequences of cyberattack on health care organizations include loss of trust, credibility, and confidence from stakeholders; in addition, the financial impact on their organization and the hospital may face regulatory sanctions [9,20,21] if due care and due process were not followed. Health care issues emanating from cyberattacks can also undermine a nation's health care policy as a whole, as the unavailability of health care systems could undermine the rights of citizens to health care [14,22].

In addition, laws have been enacted in various jurisdictions to protect the privacy of people in their countries [18,20,23]. However, data breaches in health care disrupt all these measures. According to the forecast of the International Organization for Standardization (ISO), the estimated annual losses from cybercrime could soon reach USD2 trillion [14] with countless daily breaches [19]. This forecast is in resonance with the current trend of the cost of data breaches of which health care is in the lead [7].

In this light, the European Union (EU) classified health care as an essential service having foreseen cyberattack on health care as a threat to national security [22]. This requires member states and the European Economic Area-affiliated member states to develop a culture of security across services that are vital for the economy and society and rely heavily on information and communication technology (ICT).

To maintain security in health care, various laws exist, including regulations, directives, statutory and constitutional laws, and various information security governance measures such as policies, standards, guidelines, and best practices, called "information security requirement" in this study. These were developed to prevent information security issues in health care. Owing to various cybersecurity issues, various efforts have been made to measure the security practices of health care staff [3,16,24-29], as they are the weakest link in the security chain [30,31]. However, these activities require a benchmark in the context of legal requirements in information security in health care that can be used as the measuring standard in such studies. For example, to create a questionnaire to measure health care staff's cybersecurity practices, the content of the questionnaire could be derived from the legal requirements. Therefore, the question is, what is the benchmark that is to be used as a yardstick for measuring the security compliance level of health

care staff and to what extent have these security requirements been incorporated at the organizational level where these security requirements are to be followed?

Security violations in health care facilities are not due to a lack of rule-based requirements but due to a lack of compliance with rules and in some cases due to technical vulnerabilities that could not be addressed by law, requiring an investigation as to why the challenges exist in complying with these rules. In measuring the cybersecurity practices of the health care staff, a comprehensive security requirement is required. However, a noncomprehensive security requirement is sometimes relied on, which does not serve as an effective baseline. For instance, in a recent assessment of the security practice of health care in Norway [32], the study relied on the Health Register Act, the Health Personnel Act, the Patient Records Act, and the General Data Protection Regulation (GDPR). The study relied on some legal sources; however, other vital legal sources such as the Personal Data Act of Norway, the Network and Information Security Directive of EU, and the Medical Device Directive of EU, were not considered. Other related studies [33,34] have considered a legal requirement in their work, but no study has comprehensively and systematically conducted a study on legal requirements that can serve as a benchmark for assessing health care staff security practices.

The general objective of this study is therefore to address this gap by comprehensively identifying the required security requirements in health care through state-of-the-art studies to provide input for the development of a framework for analyzing health care security practice in the context of legal requirements. The remaining sections include background studies and a specification of the scope, contribution, and research questions. This is followed by the research methods, findings, and discussion of results. A framework for analyzing health care security practice in the context of legal requirements is then presented for real studies in the future.

The health care information of persons is one of the most sensitive personal information and therefore has special protection from various laws [14,23,35,36]. Laws are rules elected to be followed by members of a society to meet the needs of society while balancing individual rights to their self-determination [37]. Laws frown against certain behaviors and are enforced by a state or the governing body. Therefore, all categories of health care information system users are legally bound to comply with legal requirements of which a contrary act will attract the application of punitive measures [20,36,38]. Therefore, it is extremely important to consider legal requirements as the baseline in measuring the security practices of health care staff.

Owing to the numerous threats of attack in health care [1-6], there have been initiatives to measure the security practices of health care staff [16]. This is to help identify the security requirements that are not being complied with and further determine the challenges or reasons why these security measures are not being complied with. The results of this study will help in finding effective solutions to enhance the conscious care behavior of users. Security practice in this study refers to how users respond to or comply with security measures that have

been established to meet the CIA requirement of systems and resources [16,24,26].

In assessing the security practices in health care, it is important to establish the scope of the hospital's legal and ethical obligations in relation to information security and privacy management [16,24,37]. This requires a catalog of comprehensive security requirements to understand the state-of-the-art legal requirements, including regulations, directives, policies, and guidelines for the fortification of users in health care IT systems against cyberattacks.

A comprehensive state-of-the-art security requirement is needed [39,40]; otherwise, what will be the benchmark in assessing the security practice level of hospital users? Moreover, if there is a security breach in health care by a user based on a lack of knowledge of a security requirement, the organization can still be liable or legally responsible [41]. This means that the health care organization will continue to make restitution for related harm caused in the breach [41]. This calls for due care and due diligence [42,43] on the part of health care organizations. Due care is measures taken by an organization to ensure that all employees are aware of acceptable and nonacceptable security practices, whereas due diligence is reasonable measures that are taken by the organizations or people to meet the established security requirements imposed by law [37]. Health care organizations increase their risk of being liable if they fail to adopt due care and due diligence measures. This is necessary because health care tends to rely more on IT and the internet for efficiency; a larger number of people can be adversely affected in a security breach situation as internet-based solutions are globally reached, which therefore require security due diligence and due care [37,42,43].

Type of Laws

Laws can be categorized based on their origins, such as constitutional law, statutory law, regulatory or administrative law, and common law, which is otherwise known as case law or precedents [37,44,45]. Constitutional law originates from the constitution of a state, bylaws, or a charter, but laws that originate from the legislative arm of governance with the mandate to make and publish laws of the country are known as statutory laws [37,44]. Furthermore, regulatory or administrative laws are created from the executive arm of the government or an authorized regulatory agency backed with executive orders and regulations [37,44]. Laws made from the judicial branch and boards based on the interpretation of law through the previous ruling of a higher court or boards are referred to as common law, case law, or precedents.

Statutory law can be further categorized into civil law and criminal law based on their association with individuals, groups, and the state [46]. Civil law has to do with issues between and among individuals and organizations [37,44] and includes contract law, employment law, and tort law. Tort law enables individuals to settle their issues in court on personal, physical, or financial matters. In such matters, restitution is settled in civil courts without the state's involvement. At the same time, criminal law is enforced and prosecuted by the state and deals with violations that are harmful to society. In criminal law, the state acts on behalf of the plaintiff to obtain retribution for the

plaintiff. For instance, in some jurisdictions, health care professionals are punished for criminal behavior if they disclose their clients' information without good causes [47].

Security Policies, Standards, Guidelines, Procedures, and Practices

In controlling information security in a health care organization, information security governance is usually adopted by organizations that use policies, standards, guidelines, procedures, and practices [37]. In various health care units, organizational policies function as the laws. Therefore, information security policies are required to be made and implemented to ensure that they are complete and appropriate and should be able to fairly apply to everyone in the workplace [37]. As laws, organizational policies must be completed with retributions, judicial practices, and sanctions to require compliance.

However, the variance between law and policy is that although ignorance of state law is not an excuse, ignorance of an organizational policy is an acceptable defense [37]. Therefore, to have an enforceable policy in an organization, the policy must be disseminated, reviewed, comprehended, complied with, and uniformly enforceable to all staff in the organization.

Information security policy directs how issues should be addressed and how IT resources should be used, but it does not define the proper operation or functioning of the system. How a software program should function is specified in the standard procedures and practices of the users' manuals and systems documentation.

Policies specify acceptable and unacceptable information security practices at the organizational level and outline rules with the aim of protecting the organization's information assets [48,49]. There are 3 types of information security policies [37,48,49]: the enterprise or organizational information security policy (EISP), issue-specific security policy (ISSP), and system-specific policy.

EISP is a general information security policy that contains the overall strategic direction, scope, and goal of the organizational information needs at a high level. In addition, EISP defines the legal requirements, outlines the responsibilities of the system administration of information security policy maintenance and practices, and outlines the responsibilities of the users.

While EISP is aimed toward addressing a broad scope of the entire organization's security issues, ISSP provides detailed guidelines pertaining to the use of specific resources, such as processor or technology, for all members or users to comply with [37,48,49]. Some of these instances include email use, internet use, security measures against viruses, bringing your own devices, use of cloud computing, home use of company-owned devices, data retention policy, and media disposal policy.

EISP and ISSP still provide information security rules at a more general level when focusing on specific systems in the organization, and they do not address security issues concerning specific systems. This gap has been filled by system-specific policy, which provides adequate information or direction in complying with the security of specific systems in the

organization [37,48-50]. System-specific policy focuses on one system such as EHR systems. In this context, system-specific policy, for instance, can be used to define the access control policy of the EHR system. Therefore, system-specific policy varies from system to system and is defined by management.

All these types of policies are effectively implemented using tools such as standards, guidelines, procedures, and practices [37,48-50]. Specifics that enable employees to comply with a security policy are known as information security standards, whereas guidelines are recommendations or examples provided to help users comply with a security policy. Practices are also recommendations or examples that are adopted from a reputable organization to help in complying with a policy, whereas procedures are step-by-step instructions users are to follow to accomplish a particular task in fulfillment of the security policy.

Scope, Contribution, and Research Questions

In assessing the information security practice of health care staff, there is a need to determine the state of security practice in the health care organization and compare it to a benchmark to determine the level of compliance with information security of the health care staff of that organization. Therefore, we opine that the legal aspect of the information security requirement is necessary to serve as the yardstick in measuring health care staff's security practices. A major reason is that a violation of any legal requirement has a huge consequence on the offending individual or company, including heavy fines, imprisonment, and payments of restitution. Therefore, aiming to comply with the legal aspect of information security requirements by using it as a yardstick will lead to unconscious compliance with the laws of that jurisdiction.

Information security requirement does not only involve legal requirements but also includes ethical security considerations of information system users [37]. However, this study focuses on the legal requirements of information security in health care such as constitutional law, statutory law, regulations, case law, and charters. Other legal sources considered in this study include information security policies and their supported instruments, such as information security standards, guidelines, and practices.

This study seeks to address issues of incomprehensiveness in considering the legal requirements for analyzing health care security practices in Norway, Ghana, and Indonesia. This has become necessary, as there have been initiatives to measure the security practices of health care staff in these countries in various projects [16]. The problem is that there is no comprehensive and state-of-the-art study of the legal requirements of information security that can serve as a baseline for assessing security practices in health care. A random and nonsystematic approach to adopting legal information security requirements in real studies could undermine the quality of the study if the baseline for the measurement is wrong. Therefore, we adopted a comprehensive, systematic scoping review approach to establish our baseline legal requirements for future imperial studies and further developed a framework to guide future related studies.

Methods

Overview

A scoping review was conducted to explore information security and privacy requirement in health care following the PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) statement [1].

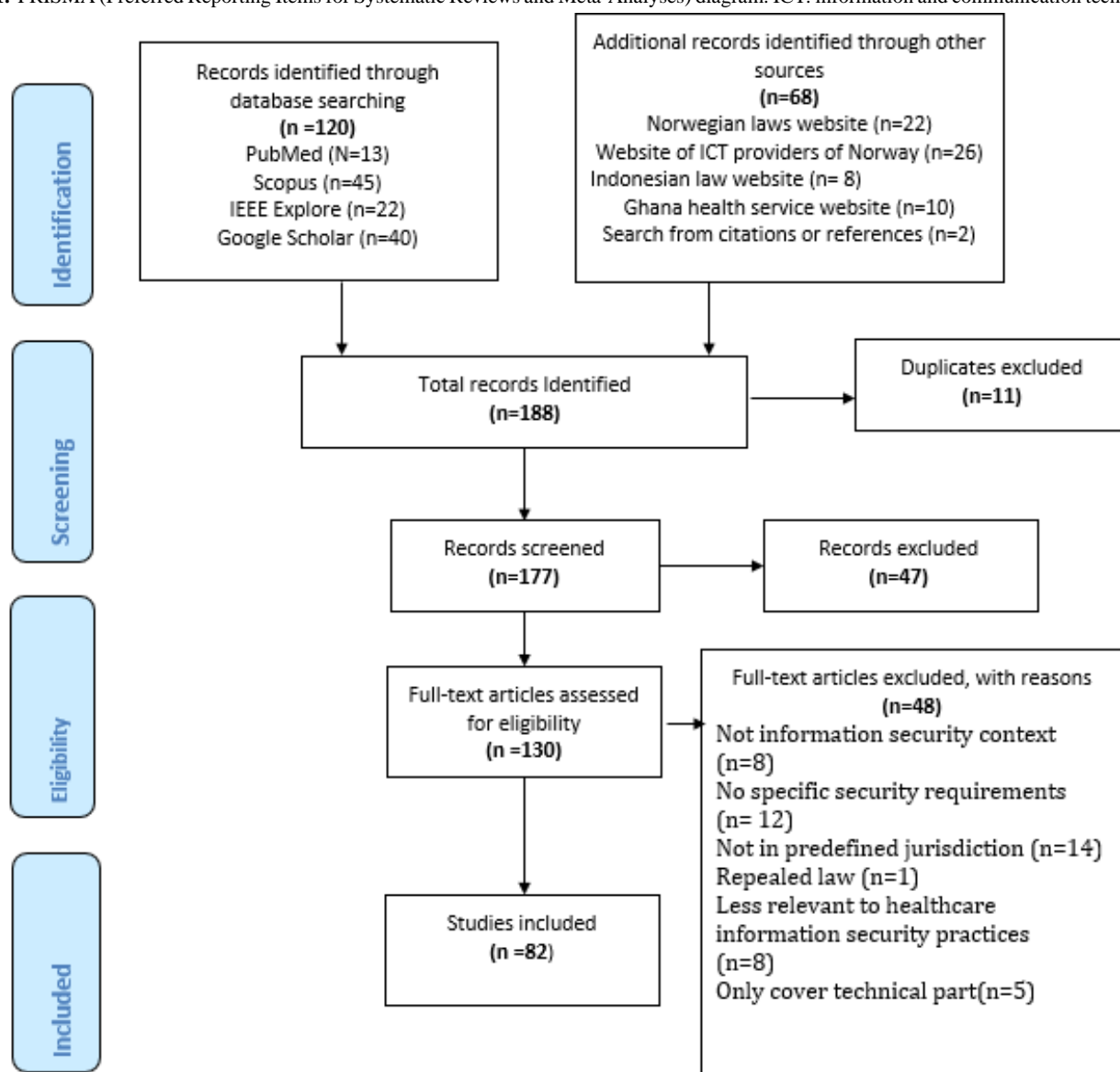
Various types of systematic studies include systematic mapping, scoping, and systematic literature review [51-54]. Systematic mapping studies rely on general research questions aimed at determining research trends or state-of-the-art studies as opposed to a scoping method that is based on the categorization of the study into topics [51,52], whereas systematic literature review aims to accumulate data with more specific research focus and synthesis. Therefore, in this study, a systematic scoping study was adapted. This section describes the methods and designs that were used to review the literature and conduct this study.

Search Strategy

The goal of the search is to search broadly to obtain comprehensive laws or rules termed here as *security requirements*. Therefore, we did not want to limit the identification of these requirements by searching through only scientifically published papers. This led to the inclusion of both scientific studies and other sources, shown in Figure 1. Therefore, the inclusion of scientific studies was intended to extract relevant laws. The sources of the security requirement were identified by conducting a literature search through several databases as follows: PubMed, Google Scholar, IEEE Xplore, and Scopus.

While reading the articles to identify the legal requirement, other relevant articles which were cited or referenced were also added in the studies and accounted for on the PRISMA diagram as *search from citations or references* as shown in Figure 1.

Figure 1. PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) diagram. ICT: information and communication technology.



In addition, we also performed manual searching through several law databases by reading all the laws under the health care category and selecting the relevant ones. The databases used were as follows:

- Legal, regulations, and directive databases for EU and Norway [55]
- Legal, regulations, directive, policy, and code of conduct databases for hospitals in Norway [55]
- Legal, regulations, and directive databases for Indonesia [56]
- Legal, regulations, directive, policy, and code of conduct databases for hospitals in Indonesia [57]
- Legal, regulations, and directive databases for Ghana [58]
- Legal, regulations, directive, policy, and code of conduct databases for hospitals in Ghana [59,60]

The literature search was conducted without time restrictions. For searching the scientific paper databases, we used the following keywords in the search string: (*Information security OR Cyber security OR Computer security*) AND *Healthcare AND Information system AND (law OR Regulation OR Directive OR Policy OR Standard)* AND (*European Union OR Norway OR Indonesia OR Ghana*). Meanwhile, for searching through law databases, we did not use any keywords. Instead, we read all the laws under the health care category and selected the relevant ones. The literature search was conducted from December 2020 to February 2021.

Eligibility Criteria

All studies involving laws, regulations, directives, policies, best practices, and standards in the health care security and privacy context in Norway and EU, Indonesia, or Ghana were eligible for review. The publication language was limited to English. Papers that did not meet the eligibility criteria or only described the technical part of security and privacy in health care without relying on legal or security governance requirements were excluded from the review. Only studies that describe the legal aspect of health care security and privacy in Norway and EU, Indonesia, or Ghana were eligible for review. Owing to the lack of resources, we focused on English scientific papers but only translated the identified local laws, which were relatively few.

Study Selection and Data Extraction

A PRISMA flow diagram of the literature search process is shown in [Figure 1](#). The titles and abstracts of articles from the databases were screened for eligibility. Then, all articles that passed the first screening entered full-text screening and data extraction. Data extraction was performed using a predesigned data collection form. For each qualified article, data on study characteristics, such as the first author and publication year, were extracted. Furthermore, we extracted information consisting of the article information, name and type of the legal document, legal document authority, security requirement, privacy requirement, health care user category, domain, responsibility level, security, and privacy requirement, which is referred to in this study as data categorization, as shown in [Table 1](#).

Table 1. Data extraction field description.

No	Category	Description
1	Paper information	Name, authors, and publication year of the paper
2	Legal document name	The name of the legal documents found in the paper
3	Legal document type	This defines the category of law such as regulation, constitutional law, directive, statutory law, policy, and guidelines found in the paper
4	Legal document jurisdiction	The country in which the legal document applies
5	Security requirement	The requirement about information security found in the legal document
6	Privacy requirement	These are the measures or rules that seek to protect the dignity of patients. These include the right to consent and the right to be forgotten to preserve the privacy of an individual
7	Health care user category	The category of users with the primary responsibility to implement or comply with the related requirement. These include management, end users, and all users. The management category includes top management such as CEOs ^a , directors, managers, and officers with the responsibility of implementing and complying with the privacy and security requirement
8	Responsibility level	The user level is responsible for the requirement, and this defines the type of user category who is to take action to observe, enforce, implement, or comply with the security measure. Examples include management, end users, and all users. The management includes top-level staff such as the CEOs, directors, managers, and officers who are responsible for implementing and observing health care security practices. End users include all employees, consultants, suppliers, and others with access to the health system. All user-level categories include responsibilities that are concerned by management and end users
9	Security category	This refers to the security domain (eg, access control, security governance, access logs, and encryption) of the requirement
10	Privacy category	This refers to the privacy domain, such as consent and right to privacy, of the requirement and data protection

^aCEO: chief executive officer.

Data Categorization

Data categorization was developed based on the objective and thorough literature reviews and author discussions. The categories were defined exclusively to assess, analyze, and evaluate the study, as shown in [Table 1](#).

Literature Evaluation

After data extraction, all researchers independently checked the extracted data. A discussion between all researchers was held to resolve all discrepancies. The selected articles were assessed, analyzed, and evaluated based on the defined categories in [Table 1](#) to evaluate the state-of-the-art security and privacy requirements. The percentages of the attributes of the categories were calculated based on the total number of counts (n) of each type of attribute. Some studies used multiple categories; therefore, the number of counts of these categories exceeded the total number of articles on the requirements presented in the study.

After data extraction, all researchers independently checked the extracted data. A discussion among all researchers was held to resolve any discrepancies.

Results

Study Selection

A total of 188 articles were identified through the literature search of the 10 databases. After duplicate deletion, 94.1% (177/188) of the articles remained for the next step. Titles and abstracts screening yielded in the exclusion of 26.6% (47/177) of the articles for not meeting eligibility criteria. Hence, 73.4% (130/177) of the articles entered the full-text screening for eligibility. After the second screening, 36.9% (48/130) of the articles were eliminated from the review for various reasons, with the main reasons being not in predefined jurisdictions

(14/48, 29%) and not having specific information security and privacy requirements (12/48, 25%). To retrieve the list of excluded papers, a request can be sent to the authors. Finally, of the 130 articles in the full-text reading stage, 82 (63.1%) met the eligibility criteria and were included for review, as shown in [Figure 1](#).

Study Characteristics

Of the 82 articles, 36 (44%) were scientific studies and the others were legal documents. A total of 75 unique legal documents were identified, including case law (n=1, 1%), charter (n=1, 1%), code of conduct (n=1, 1%), directives (n=7, 9%), guidelines (n=4, 5%), policies (n=27, 36%), recommendation (n=1, 1%), regulations (n=13, 17%), standards (n=4, 5%), and statutory law (n=16, 21%), as shown in [Multimedia Appendix 1](#) and [Table 2](#). The distribution of law jurisdictions is depicted in [Multimedia Appendix 2](#) and [Table 3](#). Of the 75 legal documents, 35 (47%) are from Norway, 9 (12%) from Ghana, 11 (15%) from Indonesia, and 17 (23%) from the EU and 3 (4%) are international laws, as presented in [Table 4](#), [Table 5](#), [Table 6](#), [Table 7](#), and [Table 8](#), respectively. In total, 253 requirements were extracted from the legal documents, consisting of 173 (68.4%) security requirements and 80 (31.6%) privacy requirements, as shown in [Multimedia Appendix 3](#). As shown in [Multimedia Appendix 4](#), of the 173 security requirements, 143 (82.7%) are the management's responsibility to fulfill, 1 (0.6%) is the end users' responsibility, and 29 (16.8%) are all users' (management and end users) responsibility. Meanwhile, as shown in [Multimedia Appendix 4](#), of the 80 privacy requirements, 70 (88%) need to be fulfilled by the management, 1 (1%) is the end users' responsibility, and 9 (11%) are all users' responsibility. Legal requirements are shown in [Table 9](#); in addition, we classified the requirements into several categories, as shown in [Tables 10](#) and [11](#).

Table 2. Types of laws (n=75).

No	Type of law	Count, n (%)
1	Case law	1 (1)
2	Charter	1 (1)
3	Code of conduct	1 (1)
4	Directive	7 (9)
5	Guideline	4 (5)
6	Policy	27 (36)
7	Recommendation	1 (1)
8	Regulation	13 (17)
9	Standard	4 (5)
10	Statutory law	16 (21)

Table 3. Count of laws based on jurisdiction (n=75).

No	Country	Count of laws, n (%)
1	Norway	35 (47)
2	Ghana	9 (12)
11	Indonesia	11 (15)
4	European Union	17 (23)
5	International	3 (4)

Table 4. Legal documents from Norway.

No	Legal document	Type
1	Code of conduct for information security and data protection in the health care and care services sector version 6.0 [61]	Code of conduct
2	Ministry of Government Administration, Reform and Church Affairs' requirements specification for PKI ^a for the public sector [62]	Guidelines
3	General principle to regional control system for information security and privacy [63]	Policy
4	Safety regulator legislation applicable to the enterprise group [63]	Policy
5	Organization of information security work [63]	Policy
6	Safety goals and level for acceptable risk of information security [63]	Policy
7	Security strategy [63]	Policy
8	Security instructions (signed version) [63]	Policy
9	ICT ^b services and information security for medical devices [63]	Policy
10	Requirements specification—ICT services and information security for MTU ^c [63]	Policy
11	Security principles and requirements for ICT infrastructure and applications [63]	Policy
12	Anonymization of health and personal information [63]	Policy
13	Use of data processor—treatment of personal information at other legal entity [63]	Policy
14	Use of email and fax [63]	Policy
15	Use of mobile phones [63]	Policy
16	Basis for posting in journal [63]	Policy
17	Storage, archiving, and deletion of health and personal information [63]	Policy
18	Crypto policy [63]	Policy
19	Password policy for the health trusts in Health South-East	Policy
20	Guidance for approval of data processing from secure third countries [63]	Policy
21	Requirements for coded research data	Policy
22	Use of email, fax, and SMS text messaging for communication with and about patients [63]	Policy
23	Regional policy for publishing and public services and DMZ ^d [63]	Policy
24	Description of identification procedure in Health South-East [63]	Policy
25	Use of logs for administrative purposes	Policy
26	Internal control information security [63]	Policy
27	Logging of activity and control of logs [63]	Policy
28	Regional security policy for cloud services [63]	Policy
29	Regulations relating to the Processing of Personal Data [64]	Regulation
30	Norwegian Personal Health Data Filing System Act [16,65,66]	Statutory law
31	Act relating to Patients' Rights	Statutory law
32	Act relating to the Processing of Personal Data [18]	Statutory law
33	Health Care Personnel Act [67,68]	Statutory law
34	Health Research Act [16]	Statutory law
35	Act relating to Public Supervision of the Health Service	Statutory law

^aPKI: public key infrastructure.

^bICT: information and communication technology.

^cMTU: medical technical equipment.

^dDMZ: demilitarized zone.

Table 5. Legal documents from Ghana.

No	Legal document	Type
1	The GHS ^a Patient's Charter	Charter
2	The Medical Profession Regulation and the Infectious Diseases, Cap 78	Regulation
3	The Ghana National Health Insurance Regulations of 2004	Regulation
4	Data Protection Act of Ghana 843	Statutory law
5	The Republic of Ghana's Constitution	Statutory law
6	The National Identification Authority Act 707	Statutory law
7	Cybersecurity Act of Ghana 2020	Statutory law
8	Guidelines for the Use of CCTV ^b in GHS Facilities	Guidelines
9	Health sector ICT ^c policy and strategy	Policy

^aGHS: Ghana Health Services.

^bCCTV: closed-circuit television.

^cICT: information and communication technology.

Table 6. Legal documents from Indonesia.

No	Legal document	Type
1	Regulation of the Minister of Health of the Republic of Indonesia Number 269/2008 on Medical Record	Regulation
2	Undang-Undang Republik Indonesia Nomor 29 Tahun 2004 Tentang Praktik Kedokteran	Statutory law
3	Undang-Undang No. 36/2009 Pasal 103 ayat 1	Statutory law
4	Peraturan Menteri Kesehatan Republik Indonesia Nomor 55 Tahun 2013 Tentang Penyelenggaraan Pekerjaan Perekam Medis	Regulation
5	Undang-Undang Republik Indonesia No 36 Tahun 2014 Tentang Tenaga Kesehatan	Statutory law
6	Peraturan Pemerintah Republik Indonesia Nomor 46 Tahun 2014 Tentang Sistem Informasi Kesehatan	Regulation
7	UU 36 Tahun 2009 Tentang Kesehatan	Statutory law
8	Peraturan Menteri Kesehatan Republik Indonesia Nomor 36 Tahun 2012 Tentang Rahasia Kedokteran	Regulation
9	Undang-Undang Republik Indonesia Nomor 44 Tahun 2009 Tentang Rumah Sakit	Statutory law
10	Peraturan Menteri Kesehatan Republik Indonesia Nomor 82 Tahun 2013 Tentang Sistem Informasi Manajemen Rumah Sakit	Regulation
11	Peraturan Menteri Kesehatan Republik Indonesia Nomor 77 Tahun 2016 Tentang Sistem Klasifikasi Keamanan Dan Akses Arsip Dinamis Di Lingkungan Kementerian Kesehatan	Regulation

Table 7. Legal documents from the EU^a.

No	Legal document	Type
1	Penal Code [41,69]	Case law
2	Directive 95/46/EC	Directive [70,71]
3	NIS ^b Directive	Directive [72]
4	The directive on patients' rights in cross-border health care (Directive 2011/24)	Directive [73]
5	Directive 2009/136/EC amending Directive 2002/58/EC (Privacy Directive)	Directive
6	Data Protection and Privacy in Electronic Communications—e-Privacy Directive (it replaces Directive 97/66/EC) [74]	Directive
7	Directive 99/93/EC	Directive [75]
8	The Patients' Rights Directive (2011/24/EU) [73]	Directive
9	Recommendation CM/Rec(2019)2 of the Committee of Ministers to member states on the protection of health-related data [76]	Guidelines
10	GCP ^c	Guidelines [71]
11	Recommendation No. R (97) 5 of the Committee of Ministers to Member States on the Protection of Medical Data	Recommendation [77]
12	GDPR ^d [16,78-83]	Regulation
13	EU regulation and compliance of national and transborder data flows	Regulation
14	Medical Device Regulation 2017/745 of EU [41]	Regulation
15	Regulation 2014/910 (the <i>eIDAS</i> ^e Regulation) [78]	Regulation
16	A European standardization group for Security and Privacy of Medical Informatics (CEN TC 251/WG6 ^f) [84,85]	Standard
17	GEHR ^g /CEN ^h standards ENV ⁱ 12265 and ENV 13606 [86,87]	Standard

^aEU: European Union.

^bNIS: Network and Information Security.

^cGCP: Good Clinical Practice.

^dGDPR: General Data Protection Regulation.

^eeIDAS: electronic identification and trust services.

^fCEN TC 251/WG6: Commission for European Normalization Technical Committee/Working Group 6.

^gGEHR: Good European Health Record.

^hCEN: European Committee for Standardization.

ⁱENV: Electronic Healthcare Record Communication for the exchange of electronic health records.

Table 8. International legal documents.

No	Legal document	Type
1	ISO ^a 27001	Standard
2	IEC ^b 80001-1:2010	Standard
3	The Universal Declaration of Human Rights	Statutory law

^aISO: International Organization for Standardization.

^bIEC: International Electrotechnical Commission.

Table 9. Legal requirement used in the study.

No	Requirement	Count, n (%)	Reference
1	GDPR ^a	13 (21.67)	[16,78-82,88-94]
2	Directive 95/46/EC	10 (16.67)	[65,70,71,74,75,95-99]
3	Norwegian Personal Health Data Filing System Act	3 (5)	[16,100,101]
4	Act relating to Patients' Rights	2 (3.33)	[16,101]
5	Act relating to the Processing of Personal Data	2 (3.33)	[16,101]
6	Directive 2011/24/EU on patients' rights in cross-border health care	2 (3.33)	[73,90]
7	Health Care Personnel Act	2 (3.33)	[16,101]
8	Act relating to Public Supervision of the Health Service	1 (1.67)	[101]
9	Data protection and privacy in electronic communications—e-Privacy Directive	1 (1.67)	[75]
10	Directive 2002/58/EC	1 (1.67)	[65]
11	Directive 2009/136/EC	1 (1.67)	[74]
12	Directive 99/93/EC	1 (1.67)	[75]
13	EU regulation and compliance of national and transborder data flows	1 (1.67)	[89]
14	GEHR ^b /CEN ^c standards ENV ^d 12265 and ENV 13606	1 (1.67)	[102]
15	Good Clinical Practice	1 (1.67)	[71]
16	Health Research Act	1 (1.67)	[16]
17	IEC ^e 80001-1:2010	1 (1.67)	[97]
18	ISO ^f 27001	1 (1.67)	[89]
19	Medical Device Regulation 2017/745 of EU	1 (1.67)	[41]
20	Ministry Of Government Administration, Reform and Church affairs' Requirements specification for PKI ^g for the public sector	1 (1.67)	[65]
21	Penal Code	1 (1.67)	[41]
22	Recommendation CM/Rec(2019)2 of the Committee of Ministers to member States on the protection of health-related data	1 (1.67)	[76]
23	Recommendation No. R (97) 5 of the Committee of Ministers to Member States on the Protection of Medical Data	1 (1.67)	[77]
24	Regulation 2014/910 (the "eIDAS Regulation")	1 (1.67)	[103]
25	Regulation of the Minister of Health of the Republic of Indonesia Number 269/2008 on Medical Record	1 (1.67)	[83]
26	Regulations relating to the Processing of Personal Data	1 (1.67)	[101]
27	The Ghana Health Services Patient's Charter	1 (1.67)	[104]
28	The Ghana National Health Insurance Regulations of 2004	1 (1.67)	[104]
29	The National Identification Authority Act 707	1 (1.67)	[104]
30	The Republic of Ghana's constitution	1 (1.67)	[104]
31	The Universal Declaration of Human Rights	1 (1.67)	[104]
32	UNDANG-UNDANG No.36/2009 and Pasal 103 ayat 1	1 (1.67)	[105]
33	Undang-undang republik, Indonesia nomor 29, Tahun 2004 tentang, Praktik kedokteran	1 (1.67)	[106]

^aGDPR: General Data Protection Regulation.

^bGEHR: Good European Health Record.

^cCEN: European Committee for Standardization.

^dENV: Electronic Healthcare Record Communication for the exchange of electronic health records.

^eIEC: International Electrotechnical Commission.

^fISO: International Organization for Standardization.

^gPKI: public key infrastructure.

Table 10. Security requirement category distribution (n=173).

No	Security requirement category	Count, n (%)
1	Data processing	14 (8.1)
2	Data protection officer	14 (8.1)
3	Right of access	13 (7.5)
4	Security by design	13 (7.5)
5	Access control	12 (6.9)
6	Email processing	10 (5.8)
7	Logs	9 (5.2)
8	Password	7 (4.1)
9	Encryption	6 (3.5)
10	Health data storage	6 (3.5)
11	Mobile phone processing	4 (2.3)
12	Privacy by design	4 (2.3)
13	CIA ^a measures	3 (1.7)
14	Data controller	3 (1.7)
15	Personal data	3 (1.7)
16	Third countries	3 (1.7)
17	Data protection	3 (1.7)
18	Backup	2 (1.2)
19	Documentation	2 (1.2)
20	Electronic signature	2 (1.2)
21	Establish security governance	2 (1.2)
22	Least privileges	2 (1.2)
23	Medical devices	2 (1.2)
24	Right to be informed	2 (1.2)
25	Risk management	2 (1.2)
26	Security governance	2 (1.2)
27	Third parties	2 (1.2)
28	Data breach	2 (1.2)
29	Use of ISO ^b standards	2 (1.2)
30	Consent	1 (0.6)
31	Data aggregation	1 (0.6)
32	Incident reporting	1 (0.6)
33	Internal control	1 (0.6)
34	Data transfer to non-EU ^c countries	1 (0.6)
35	Deletion of health data	1 (0.6)
36	Establish security policies	1 (0.6)
37	Health care data hosting	1 (0.6)
38	Identity	1 (0.6)
39	Internal and external threats	1 (0.6)
40	Mobile devices	1 (0.6)
41	Monitoring of NIS ^d Directives	1 (0.6)

No	Security requirement category	Count, n (%)
42	Patients from other member states	1 (0.6)
43	Physical security	1 (0.6)
44	Professional secrecy	1 (0.6)
45	Protection against security incidents	1 (0.6)
46	Providing information to patients from a member state	1 (0.6)
47	Risk assessment	1 (0.6)
48	Risk mitigation	1 (0.6)
49	Sanction	1 (0.6)
50	Technological security measures	1 (0.6)
51	Training and education	1 (0.6)

^aCIA: confidentiality, integrity, and availability.

^bISO: International Organization for Standardization.

^cEU: European Union.

^dNIS: Network and Information Security.

Table 11. Privacy requirement category distribution (n=80).

No	Privacy requirement category	Count, n (%)
1	Consent	13 (16)
2	Disclosure of health data	12 (15)
3	Privacy by design	8 (10)
4	Right to privacy	8 (10)
5	Right of access	7 (9)
6	Data protection	6 (8)
7	Data processing	3 (4)
8	Personal data	3 (4)
9	Punitive measures of security and privacy violation	3 (4)
10	How to record health data	2 (3)
11	Privacy rights	2 (3)
12	Storage of health records	2 (3)
13	CIA ^a measures	1 (1)
14	Data collection purpose	1 (1)
15	Deletion of health data	1 (1)
16	Electronic signatures	1 (1)
17	Mobile phone processing	1 (1)
18	Professional secrecy	1 (1)
19	Purpose of health care data processing	1 (1)
20	Right to be forgotten	1 (1)
21	Right to object	1 (1)
22	Termination of consent	1 (1)
23	Third parties	1 (1)

^aCIA: confidentiality, integrity, and availability.

Findings

The following sections present and describe a series of findings, including law by type, law by jurisdiction, requirement by type, requirement by responsibility level, and identified security and privacy requirements and their categorizations.

Law by Type

The types of laws identified in this work are presented in [Multimedia Appendix 1](#) and [Table 2](#). A total of 75 legal requirements were identified in this review. The most common types of laws that were used are policies (27/75, 36%), statutory law (16/75, 21%), regulations (13/75, 17%), directive (7/75, 9%), standards (4/75, 5%), and guidelines (4/75, 5%), but recommendation, code of conduct, charter, and case law accounted for the lowest proportion. It is worth noting that the 27 policies were all collected from information security policy documents of the health care facilities of the southeast region in Norway as their internal control measures of information security and privacy measures.

Law by Jurisdiction

The specific legal documents from Norway, Ghana, Indonesia, the EU level, and the international level are listed in [Table 4](#), [Table 5](#), [Table 6](#), [Table 7](#), and [Table 8](#), respectively, and Norway has almost half (36/75, 48%) of the laws pertaining to information security and privacy, which were identified in this work and shown in [Multimedia Appendix 2](#) and [Table 4](#). This was followed by the EU (17/75, 23%). The southeast health region in Norway developed approximately 27 policies, which also accounted for the larger proportion of the laws in Norway than that in other countries, as shown in the bar chart of the law jurisdiction distribution in [Multimedia Appendix 2](#).

Identified Legal Requirement

Of the 82 requirement sources, 36 (44%) were articles that considered at least one of the identified requirements, whereas the others were legal documents. In total, 75 unique legal documents were identified, and 33 legal documents were identified to have been considered in the papers as shown in [Table 9](#).

Moreover, as shown in [Table 9](#), among all the legal documents, the GDPR (13/60, 22%) is the most common regulation that was used in the articles that relied on legal requirements, followed by Directive 95/46/EC (10/60, 17%), which has already been repealed and replaced by the GDPR. Some acts from Norway, as well as directive from the EU, have also been referred to several times, such as the Norwegian Personal Health Data Filing System Act (3/60, 5%), Act relating to Patients' Rights (2/60, 3%), Act relating to the Processing of Personal Data (2/60, 3%), Directive 2011/24/EU on patients' rights in cross-border health care (2/60, 3%), and Health Care Personnel Act (2/60, 3%).

Security and Privacy Requirements

According to [Multimedia Appendix 3](#), most legal requirements extracted are security requirements (173/253, 68.4%), whereas the rest are privacy requirements (80/253, 31.6%).

Requirements by Responsibility Level

The identified responsibility level of users includes management, end users, and all users. The management level has more security and privacy responsibility and stipulation than the end users. As shown in [Multimedia Appendices 4](#) and [5](#), documents list the security and privacy requirements only for end users.

Security Category

The security requirements extracted from all the studies cover various aspects, such as data processing, data protection officer, right of access, security by design, access control, email processing, logs, and password, as shown in [Table 10](#). In this study, security requirements relating to data processing (14/173, 8.1%), data protection officer (14/173, 8.1%), right of access (13/173, 7.5%), security by design (13/173, 7.5%), access control (12/173, 6.9%), email processing (10/173, 5.8%), logs (9/173, 5.2%), password (7/173, 4%), encryption (6/173, 3.5%), and health data storage (6/173, 3.5%) were identified to be commonly adopted in the legal requirements, as shown in [Table 10](#).

Privacy Category

The privacy requirement categories that were realized in this work are shown in [Table 11](#).

The areas that were mostly required by the legal instruments are consent (13/80, 16%), disclosure of health data (12/80, 15%), privacy by design (8/80, 10%), right to privacy (8/80, 10%), right of access (7/80, 9%), data protection (6/80, 8%), data processing (3/80, 4%) and punitive measures (3/80, 4%).

Discussion

Principal Findings

The main purpose of this study is to comprehensively identify, assess, and synthesize the appropriate legal requirements and security governance tools of information security to serve as a yardstick for modeling and analyzing health care security practices. A scoping review of these requirements was conducted to include various categories, as presented in [Table 1](#). The most used categories identified in this study are listed in [Table 12](#). For instance, among various types of laws that were identified in this study ([Multimedia Appendix 1](#)), the most used types of law are the policies, statutory law, regulations, and directives, as shown in [Table 12](#).

Table 12. Summary of the most used categories.

No	Category	Most used
1	Type of law	Policy, statutory law, regulation, and directive
2	Jurisdiction	Norway and European Union
3	Requirement type	Security requirement
4	Responsibility level	Management
5	Security requirement category	Data processing, data protection officer, right of access, security by design, access control, email processing, logs, password, encryption, and health data storage
6	Privacy requirement category	Consent, disclosure of health data, privacy by design, right of access, and data protection

Security Requirement Responsibility Level Distribution

As defined in [Table 1](#), the responsibility level of the requirement is the level of user categories that take action to observe, enforce, implement, or comply with the security measure. Examples include management, end users, and all users. Management includes top-level staff, such as the chief executive officers (CEOs), directors, managers, and officers, who are responsible for implementing and observing health care security practices. All users include all employees, consultants, suppliers, and others with access to the health care system and with the responsibility to comply with security and privacy requirements. The end users' level includes only those user categories that have access to the health care system with the purpose of accessing and performing specified tasks. Such users include nurses, doctors, pharmacies, record management, and patients' EHRs for therapeutic reasons.

As shown in [Multimedia Appendices 4 and 5](#), the management level was identified to be mostly responsible for information security and privacy requirements, followed by *all users*. This implies that in most information security and privacy requirement categories such as access control, password management, consent, and incident reporting, as outlined in [Tables 10 and Tables 11](#), the management level has more responsibility. The management user category includes the CEO, chief information officer, chief information security officer, all directors, and all managers responsible for formulating, designing, and implementing privacy and security policies for compliance [37]. The top-management user category, such as the CEO, chief information officer, and chief information security officer, is responsible for coming out with the information security governance requirement based on prevailing laws pertaining to information security. Directors and managers then ensure that the policies, guidelines, standards, and best practices are appropriately designed and implemented. They also need to create awareness and ensure that all personnel are adequately trained in these requirements. Essentially, impact assessments such as privacy and security are also conducted by the management. To ensure compliance, these policies need to be monitored and evaluated. Management, therefore, has a major proportion of responsibility because of all these broad activities being performed toward enhancing security.

In addition, the *all users* category consists of all employees such as the management level and end users including temporal workers and contractors who have the responsibility to enforce and comply with the requirements. The *all users* category of

the level of responsibility involves requirements that need the attention of both management and end users. For instance, access control requires management to incorporate it into the development of systems. However, end users must also be responsible for their access control-related behaviors, including password management. The *end users* level includes those health care workers who are given access to a system based on their need to use that system for therapeutic purposes [61]. Examples include the end users of an EHR system. This group of users is mostly large in number but does not have an enormous number of responsibilities as compared with the management group, as shown in [Multimedia Appendices 4 and 5](#).

Requirement Types (Security and Privacy)

A total of 2 kinds of measures were extracted from the legal documents in this study, namely, security and privacy requirements. The legal documents contain at least one of the two kinds of measures: privacy, security, or both. Furthermore, >1 requirement was found in some of the sources of the legal documents, and this resulted in more legal requirements compared with the number of identified sources, as shown in [Table 9](#). After the identification and extraction process, 173 security requirements and 80 privacy requirements were identified, as shown in [Multimedia Appendix 3](#). The findings indicate that there are more security requirements than privacy requirements identified in this study. The main reason is that many policies in Norway describe security requirements, as shown in [Multimedia Appendix 1 and Table 4](#). Most of these policies were developed to address security requirements such as email use, crypto policy, password policy, and access control logging, which resulted in the number of security requirements surpassing the number of privacy requirements.

Law by Type

From [Table 2](#), a total of 10 types of laws were identified in this study, including case law, charter, code of conduct, directives, guidelines, policies, and recommendations. Others include regulations, standards, and statutory law, of which the most used type of laws are policies (27/75, 36%), statutory law (16/75, 21%), regulations (13/75, 17%), directives (7/75, 9%), standards (4/75, 5%), and guidelines (4/75, 5%), as shown in [Table 12](#). The standards that were identified are only from the EU and international levels with which Norway is bound to comply. In addition, none of the countries has standards as far as what we have collected. This could be due to the level of maturity of IT development in health care in each country. Finally, only a few

documents were categorized into case law, charter, recommendation, and code of conduct.

One of the most influential legal documents that covers almost every general aspect, as mentioned is the GDPR, as shown in [Table 9](#), to which data controllers, data processors, and data subjects need to comply. It is worth mentioning that pursuant to the GDPR, “a data controller is a legal person, public authority, agency or other body which, alone or jointly with others, determines the purposes and means of the processing of personal data,” whereas a data processor means a legal person, public authority, agency, or other body that processes personal data on behalf of the controller [107]. A data subject is any identified or identifiable person whose data are processed by the data processor. ISO 27001 provides a framework for managing security issues in health care including the measures covering information security policies, organization of information security, human resource security, asset management, media handling, access control, cryptography, physical and environmental security, operational security, communications security, system acquisition, development and maintenance, supplier relationships, and information security incident management through ISO 27799 [14]. Health care has extended needs in these areas, which is why ISO 27799 was developed for use in conjunction with ISO 27001. This provides room to address the security and privacy requirements that have not been fully covered in ISO 27001.

The widely used model, namely, the CIA triad, which is the balanced protection of CIA of data [108], is the foundation and basis of many laws and regulations including the GDPR, Recommendation CM/Rec (2019)2 of the Committee of Ministers to member states on the protection of health-related data, Directive 2009/136/EC amending Directive 2002/58/EC (Privacy Directive), Medical Device Regulation 2017/745 of EU, and Regulation 2014/910 (the *eIDAS Regulation*) at the EU level, as well as the Norwegian Personal Health Data Filing System Act, Act relating to the Processing of Personal Data, and Act relating to Patients’ Rights as shown in [Table 9](#).

Law by Country

The legal documents were identified from 3 countries: Norway, Ghana, and Indonesia. Norway has the most legal documents for this study at 47% (35/75), whereas Ghana and Indonesia provide only 12% (9/75) and 15% (11/75) of the documents, respectively. The main reason Norway has far more relevant legal documents than the other 2 countries is that Norway has many policies that describe specific details on security and privacy requirements. Furthermore, we also identified some legal documents from the EU (17/75, 23%) and some international laws (3/75, 4%). Most EU documents are directives and regulations that should be adopted by EU members, including Norway. Meanwhile, the international laws include 2 ISOs and 1 statutory law, which should be adopted by all countries.

Security and Privacy Policies in Norway, Ghana, and Indonesia

The privacy requirements in this study focused on patients’ consent to the processing of their personal data and the

processing and storage of medical records, as shown in [Table 11](#). The requirements for processing personal information include that the data subjects must consent to the use of their data captured and collected in the first place [109]. Patients have the right to object to the processing of their personal health data (Norwegian Personal Health Data Filing System Act [110]) and are entitled to their information not to be disclosed to a third party without their consent [111]. The Health Research Act in Norway stipulates that more detailed requirements regarding consent must be informed, voluntary, express, and documented [112]. As for the processing of medical records, it is specifically stated in Indonesian laws that the medical data should be kept confidential by the management level to protect the patients and hospitals must protect archived physical records [106].

Security and privacy requirements in Norway, Ghana, and Indonesia all contain laws to protect the CIA of health care data. As shown in [Multimedia Appendix 2](#), almost 46% (35/75) of the laws were developed by Norway, and most of the information security and privacy policies were developed by Norwegian health care facilities to meet the CIA requirements of health care data and information, as compared with Indonesia and Ghana. The variance could arise from various reasons, including advancement in the application of ICT in health care between European and African countries [113,114], and culturally related factors among the 3 countries. Norway is one of the countries in Europe that might have been more advanced in the use of ICT in health care than Ghana and Indonesia and have therefore adopted more legal requirements than Ghana and Indonesia. In addition, Norway is affiliated with the EU through the European Economic Area and is therefore bound to adopt the legal requirements, such as the GDPR and Network and Information Security Directive. In addition, EU countries, including Norway, are concerned with privacy [114]. This may have been one of the reasons for the adoption of more legal requirements to comprehensively enhance privacy and security measures.

Framework

On the basis of our findings on security requirements, we present a framework in this section to provide directions for future imperial research in health care security practices. The framework consists of comprehensive security practices (drawn from the security requirements) and categories of health care staff in health care information security practices. It also includes analysis methods, the actual measure of security practices in a typical hospital, a gap or security failures, and an incentivization module, as shown in [Figures 2 and 3](#) and as described as follows:

- Comprehensive security requirements: these include both privacy and security requirements that have been identified in the legal and security governance requirements in this study, as shown in [Tables 10 and 11](#). These requirements are to be observed by all categories of health care workers. These requirements serve as the benchmark to be complied with by all categories of health care staff.
- Categories of users: these include management, all users, and the end users of a typical hospital. These categories of

- users must observe the required security practices at their respective levels, as shown in Figure 2.
- Analysis methods: in assessing health care security practices, various methods must be identified and used, as shown in Figure 2. These include a hybrid survey consisting of both qualitative and quantitative approaches [6,16-115]. Attack-defense simulation is when the investigator acts as the adversary to gain access to health care resources by using various techniques, including social engineering, brute-force attack, and SQL injection, depending on the goal of the attacker. Data analysis with machine learning can also be adopted to analyze logs of health care staff to determine abnormal access and maliciousness. The analysis method obtains inputs from the comprehensive required security and privacy practices fused with the various levels of health care staff user categories.
 - In addition, health care staff have various characteristics that can be traced in the psychological-social and cultural contexts, social engineering, and access logs [16].
 - These qualities also serve as input to the study approach.
 - The actual measure of security practices was then determined from the assessment and compared with the required security and privacy practices.

- Security failures are gaps or deltas in the security practices that are determined if, after assessment, the hospital is not able to fully comply with all the identified requirements.
- Security and privacy enhancement measures: security failures can be improved with security and privacy enhancement measures, such as incentive measures and improving on factors that influence security failures. For instance, health care staff can be treated with various incentivization measures to improve their security-conscious care behavior. The assessment can then be conducted to determine the effectiveness of the treatment.

Information security and privacy requirements change based on or assessed threats, thus requiring changes in various laws. Therefore, the framework is such that the study can always be repetitive, as shown in Figure 2, to assess and identify related security and privacy gaps among health care workers in their application of ICT in health care. In Figure 2, the framework implementation is simplified, and security requirements are identified for security and privacy behavior assessment. The findings were compared with the required security behavior. Identified gaps can always be improved through cybersecurity and privacy incentives.

Figure 2. Legal requirement framework.

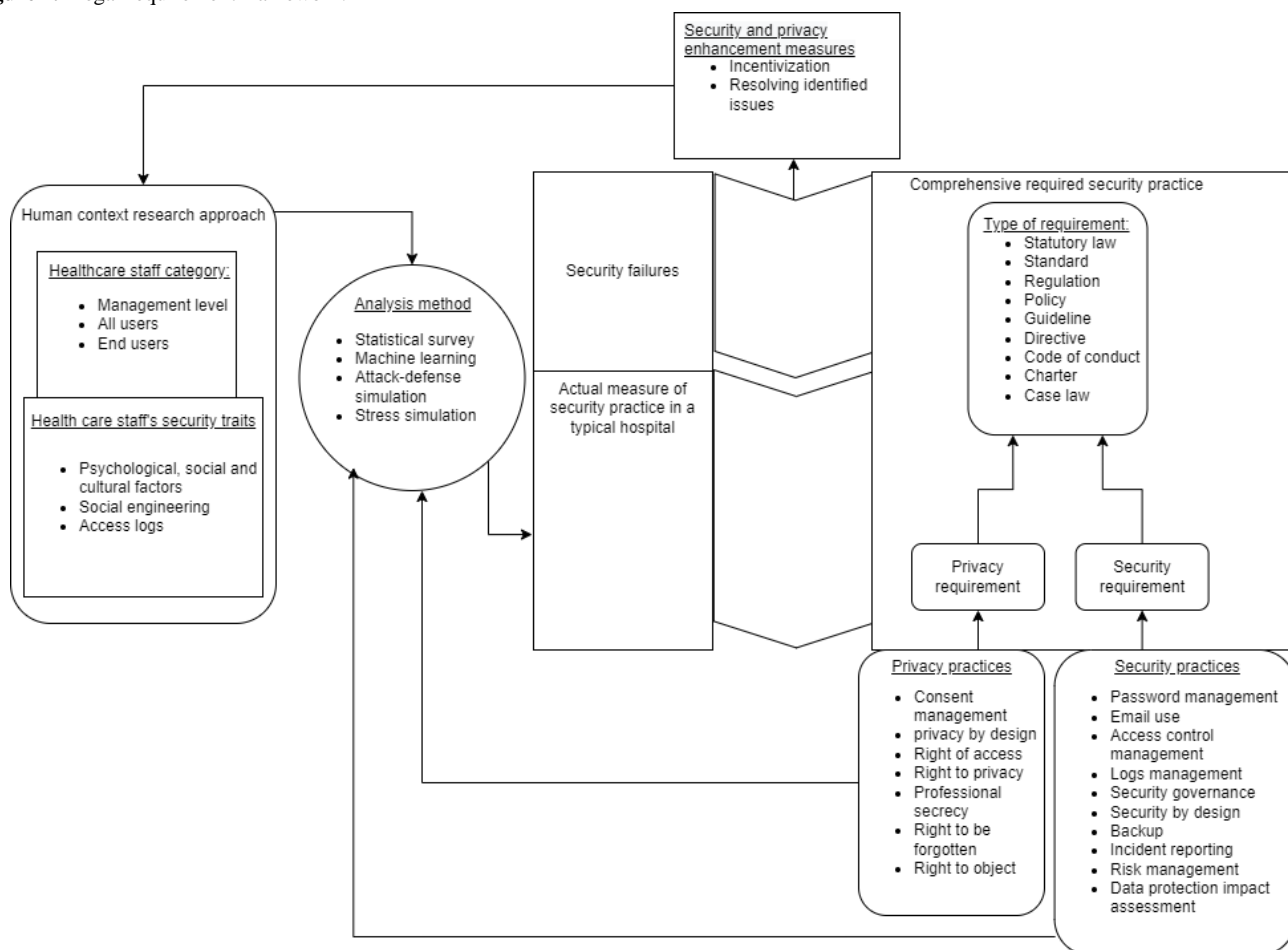
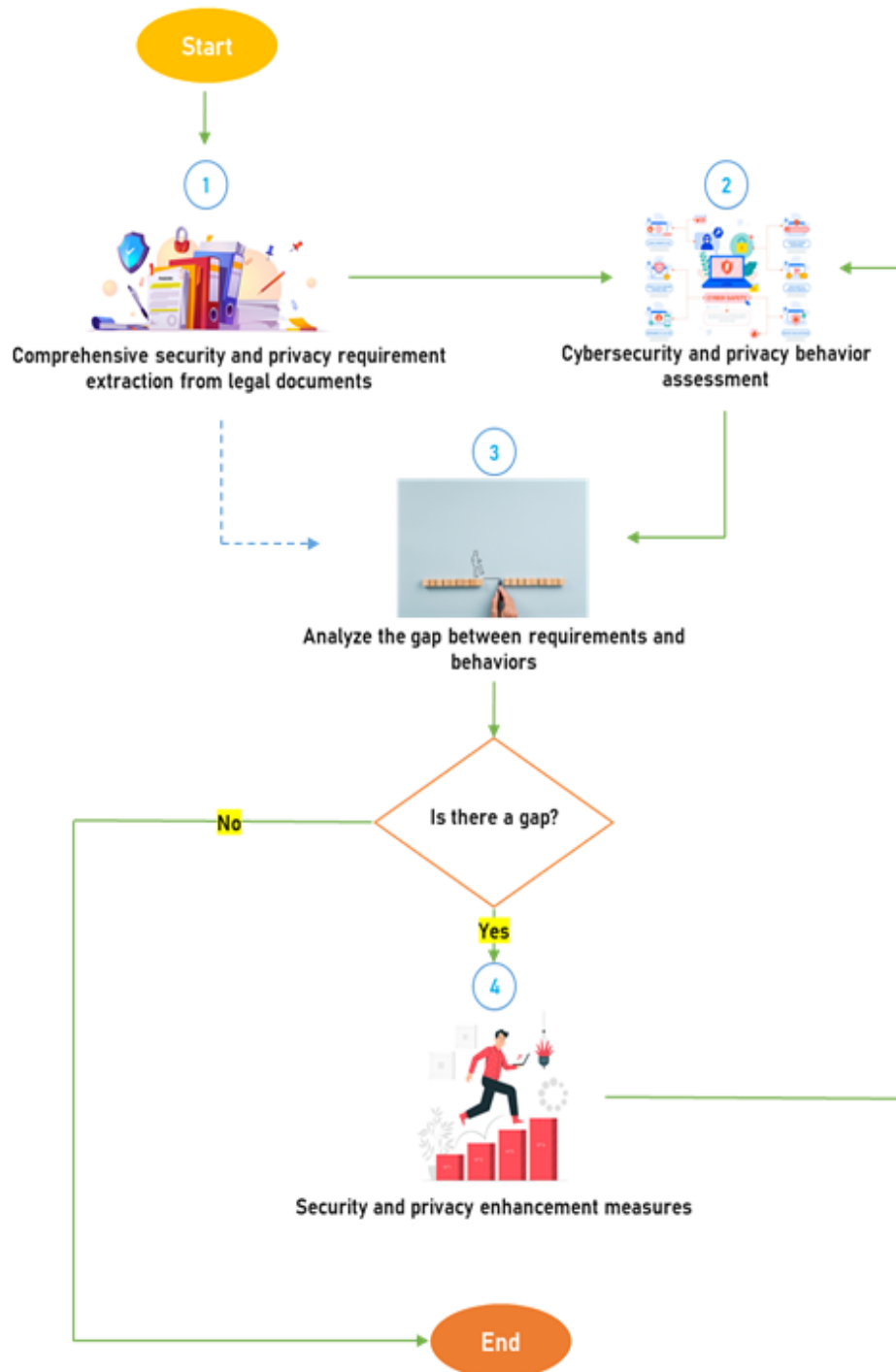


Figure 3. Measurement flowchart.



Conclusions

Amidst various information security solutions, data breaches continue to increase, especially in the area of the health care staff information security practice. This has attracted research interest in modeling and assessing health care staff's information security practices toward improving their security-conscious care behavior.

However, there is no holistic benchmark that serves as a yardstick in assessing health care information security practices comprehensively. To this end, we systematically reviewed information security requirements in health care in the context

of legal requirements and information security governance tools for comprehensive security and privacy requirements in health care in Norway, Indonesia, and Ghana. Approximately 173 security requirements covering data processing, right of access, security by design, access control, email processing, logging, password, encryption, health care data storage, data processing officer, and so on were identified, as shown in [Table 10](#).

In addition, approximately 80 privacy requirement categories were identified and included consent, disclosure of health data, privacy by design, right to privacy, right of access, data protection, data processing, personal data, and punitive measures, as shown in [Table 11](#). On the basis of these findings,

a framework for modeling, analyzing, and developing effective security countermeasures, including incentivization measures, was developed, as shown in [Figures 2](#) and [3](#). Following this framework, research results of health care security practices would be more reliable and effective than relying on incomprehensive security requirements. However, we observed some limitations that should be considered in future studies. For instance, there may be more standards in information security, but we focused on health care-related information security standards from the scientific papers that we searched for based on the scope we set. Therefore, it may not be an

exhaustive list of information security standards. Although we have identified the requirements and practices, in this framework, our work has not taken measures to narrow down the gap between requirements and practices by way of a real implementation. This is another limitation, and will be the next step in future work.

Having postulated this, the framework must be implemented to assess its effectiveness for general use. This framework will serve as a guideline for assessing security practices in health care.

Conflicts of Interest

None declared.

Multimedia Appendix 1

Law type distribution.

[[PNG File , 58 KB](#) - [humanfactors_v9i2e30050_app1.png](#)]

Multimedia Appendix 2

Law jurisdiction distribution. EU: European Union.

[[PNG File , 35 KB](#) - [humanfactors_v9i2e30050_app2.png](#)]

Multimedia Appendix 3

Requirement type distribution.

[[PNG File , 91 KB](#) - [humanfactors_v9i2e30050_app3.png](#)]

Multimedia Appendix 4

Security requirement responsibility level.

[[PNG File , 86 KB](#) - [humanfactors_v9i2e30050_app4.png](#)]

Multimedia Appendix 5

Requirement type distribution.

[[PNG File , 65 KB](#) - [humanfactors_v9i2e30050_app5.png](#)]

References

1. Widup S. 2019 Verizon Data Breach Investigations Report. 2019. URL: <https://www.nist.gov/system/files/documents/2019/10/16/1-2-dbir-widup.pdf> [accessed 2022-03-22]
2. Devis J. The 10 biggest healthcare data breaches of 2019, so far. xtelligent Healthcare Media. 2019. URL: <https://healthitsecurity.com/news/the-10-biggest-healthcare-data-breaches-of-2019-so-far> [accessed 2022-03-22]
3. Yeng PK, Nweke LO, Woldaregay AZ, Yang B, Snekenes EA. Data-driven and Artificial Intelligence (AI) approach for modelling and analyzing healthcare security practice: a systematic review. In: Intelligent Systems and Applications: Proceedings of the 2020 Intelligent Systems Conference (IntelliSys) Volume 1. Cham, Switzerland: Springer; Aug 25, 2020:1-18.
4. Attacks on health care in the context of COVID-19. World Health Organization. 2020 Jul 30. URL: <https://www.who.int/news-room/feature-stories/detail/attacks-on-health-care-in-the-context-of-covid-19> [accessed 2022-03-22]
5. Brno University Hospital ransomware attack (2020). Cyber Law Toolkit. 2020 Mar 13. URL: [https://cyberlaw.ccdcoe.org/w/index.php?title=BrnoUniversityHospitalransomwareattack\(2020\)&oldid=2290](https://cyberlaw.ccdcoe.org/w/index.php?title=BrnoUniversityHospitalransomwareattack(2020)&oldid=2290) [accessed 2022-03-22]
6. Muthuppalaniappan M, Stevenson K. Healthcare cyber-attacks and the COVID-19 pandemic: an urgent threat to global health. *Int J Qual Health Care* 2021 Feb 20;33(1):117 [FREE Full text] [doi: [10.1093/intqhc/mzaa117](https://doi.org/10.1093/intqhc/mzaa117)] [Medline: [33351134](https://pubmed.ncbi.nlm.nih.gov/33351134/)]
7. Cost of a Data Breach Report 2020. IBM. 2020. URL: <https://www.ibm.com/account/reg/us-en/signup?formid=urx-46542> [accessed 2022-03-22]
8. Ayyagari R. An exploratory analysis of data breaches from 2005-2011: trends and insights. *J Inf Priv Secur* 2014 Jul 07;8(2):33-56. [doi: [10.1080/15536548.2012.10845654](https://doi.org/10.1080/15536548.2012.10845654)]

9. Moffit RE, Steffen B. Health Care Data Breaches: A Changing Landscape. Maryland Health Care Commission. 2017 Jun. URL: https://mhcc.maryland.gov/mhcc/pages/hit/hit/documents/HIT_DataBreachesBrief_Brf_Rpt_090717.pdf [accessed 2022-03-22]
10. Hjellen B. Hacking scandal shakes Finland-patients pressured for money. NRK. 2020 Oct 25. URL: <https://www.nrk.no/urix/hacking-skandale-ryster-finland---pasienter-presset-for-penger-1.15214710> [accessed 2022-03-22]
11. A German hospital hacked, patient taken to another city dies. Associated Press. 2020 Sep 17. URL: <https://www.securityweek.com/german-hospital-hacked-patient-taken-another-city-dies> [accessed 2022-03-22]
12. Humer C, Finkle J. Your medical record is worth more to hackers than your credit card. Reuters. 2014 Sep 24. URL: <https://www.reuters.com/article/us-cybersecurity-hospitals-idUSKCN0HJ21I20140924> [accessed 2022-01-22]
13. The Code of Conduct for information security and data protection in the healthcare and care services is a holistic approach to an information security policy for all organisations within the sector. Direktoratet for e-helse. URL: <https://www.ehelse.no/normen/documents-in-english> [accessed 2022-02-22]
14. Health informatics-Information security management in health using ISO/IEC 27002. International Organization for Standardization. 2016. URL: <https://www.iso.org/standard/62777.html> [accessed 2022-01-22]
15. Canny SD, Salam AF. A framework for health care information assurance policy and compliance. Commun ACM 2010 Mar;53(3):126-131. [doi: [10.1145/1666420.1666453](https://doi.org/10.1145/1666420.1666453)]
16. Yeng PK, Yang B, Snekenes EA. Framework for healthcare security practice analysis, modeling, and incentivization. In: Proceedings of the 2019 IEEE International Conference on Big Data. 2019 Presented at: BigData '19; December 9-12, 2019; Los Angeles, CA, USA p. 3242-3251. [doi: [10.1109/BigData47090.2019.9006529](https://doi.org/10.1109/BigData47090.2019.9006529)]
17. Patil HK, Seshadri R. Big data security and privacy issues in healthcare. In: Proceedings of the 2014 IEEE International Congress on Big Data. 2014 Presented at: BigData '14; June 27-July 2, 2014; Anchorage, AK, USA p. 762-765. [doi: [10.1109/bigdata.congress.2014.112](https://doi.org/10.1109/bigdata.congress.2014.112)]
18. Proposition 56 LS (2017-2018)/Act on the processing of personal data (the Personal Data Act). Ministry of Justice and Public Security. 2017. URL: <https://www.regjeringen.no/no/dokumenter/prop-56-ls-/id2594627/> [accessed 2022-03-22]
19. Lewis B. How to tackle today's IT security risks. International Organization for Standardization. 2019. URL: <https://www.iso.org/news/ref2360.html> [accessed 2022-03-22]
20. Nadeau M. General Data Protection Regulation (GDPR): What you need to know to stay compliant. CSO India. 2020 Jun 12. URL: <https://www.csoonline.com/article/3202771/general-data-protection-regulation-gdpr-requirements-deadlines-and-facts> [accessed 2022-12-01]
21. Key Changes with the General Data Protection Regulation. General Data Protection Regulation. 2019. URL: <https://eugdpr.org/the-regulation/> [accessed 2022-03-22]
22. Directive (EU) 2016/1148 of the European Parliament and of the Council of 6 July 2016 concerning measures for a high common level of security of network and information systems across the Union. European Commission. 2017. URL: <https://eur-lex.europa.eu/legal-content/EN/ALL/?uri=CELEX%3A32016L1148> [accessed 2022-03-22]
23. Ghana Data Protection Act, 2012 (ACT 843). Data Protection Commission. 2012. URL: <https://www.dataprotection.org.gh/index.php/data-protection/data-protection-acts-2012> [accessed 2021-12-22]
24. Yeng PK, Yang B, Snekenes EA. Observational measures for effective profiling of healthcare staffs' security practices. In: Proceedings of the 43rd Annual Computer Software and Applications Conference. 2019 Presented at: COMPSAC '19; July 15-19, 2019; Milwaukee, WI, USA p. 397-404. [doi: [10.1109/COMPSAC.2019.10239](https://doi.org/10.1109/COMPSAC.2019.10239)]
25. Nweke LO, Yeng PK, Wolthusen SD, Yang B. Understanding attribute-based access control for modelling and analysing healthcare professionals' security practices. Int J Adv Comput Sci Appl 2020;11(2):683-690. [doi: [10.14569/ijacsa.2020.0110286](https://doi.org/10.14569/ijacsa.2020.0110286)]
26. Yeng PK, Yang B, Snekenes EA. Healthcare staffs' information security practices towards mitigating data breaches: A literature survey. Stud Health Technol Inform 2019;261:239-245. [Medline: [31156123](https://pubmed.ncbi.nlm.nih.gov/31156123/)]
27. Kim J, Grillo JM, Boxwala AA, Jiang X, Mandelbaum RB, Patel BA, et al. Anomaly and signature filtering improve classifier performance for detection of suspicious access to EHRs. AMIA Annu Symp Proc 2011;2011:723-731 [FREE Full text] [Medline: [22195129](https://pubmed.ncbi.nlm.nih.gov/22195129/)]
28. Box D, Pottas D. Improving information security behaviour in the healthcare context. Proc Technol 2013;9:1093-1103. [doi: [10.1016/j.protcy.2013.12.122](https://doi.org/10.1016/j.protcy.2013.12.122)]
29. Box D, Pottas D. A model for information security compliant behaviour in the healthcare context. Proc Technol 2014;16:1462-1470. [doi: [10.1016/j.protcy.2014.10.166](https://doi.org/10.1016/j.protcy.2014.10.166)]
30. How to preventing social engineering attacks protect your business data. RMON Networks. 2016 Oct 20. URL: <https://rmonnetworks.com/prevent-social-engineering-phishing/> [accessed 2019-12-22]
31. Arce I. The weakest link revisited [information security]. IEEE Secur Privacy 2003 Mar;1(2):72-76. [doi: [10.1109/msecp.2003.1193216](https://doi.org/10.1109/msecp.2003.1193216)]
32. Assessment of the health trusts' prevention of attacks on their ICT systems-Office of the Auditor General of Norway annual report. Riksrevisjonen. 2020. URL: <https://www.riksrevisjonen.no/contentassets/1855de07d48f4292a8d9b28fd8795af6/annualreport2020.pdf> [accessed 2022-05-22]

33. Jensen J, Tøndel IA, Jaatun MG, Meland PH, Andresen H. Reusable security requirements for healthcare applications. In: Proceedings of the 2009 International Conference on Availability, Reliability and Security. 2009 Presented at: ARES '09; March 16-19, 2009; Fukuoka, Japan p. 380-385. [doi: [10.1109/ares.2009.107](https://doi.org/10.1109/ares.2009.107)]
34. Power DJ, Politou EA, Slaymaker MA, Simpson AC. Towards secure Grid-enabled healthcare. *Softw Pract Exper* 2005 Jul 25;35(9):857-871. [doi: [10.1002/spe.692](https://doi.org/10.1002/spe.692)]
35. Hartvigsen G, Pedersen S. Lessons learned from 25 years with telemedicine in Northern Norway. University Hospital of North Norway, Norwegian Centre for Integrated Care. 2015 Jun 23. URL: <https://www.isftech.org/files/media/TitulPrefaceContent.pdf> [accessed 2022-04-12]
36. Implementation of GDPR in health care sector in Norway. Direktoratet for e-helse. 2019. URL: <https://tinyurl.com/yckn3srn> [accessed 2022-03-22]
37. Whitman ME, Mattord HJ. Principles of Information Security. 4th edition. Boston, MA, USA: Cengage Learning; 2011.
38. EU GDPR Information Portal. General Data Protection Regulation. 2018. URL: <http://eugdpr.org/html> [accessed 2021-03-22]
39. Nasiri S, Sadoughi F, Tadayon MH, Dehnad A. Security requirements of Internet of things-based healthcare system: a survey study. *Acta Inform Med* 2019 Dec;27(4):253-258 [FREE Full text] [doi: [10.5455/aim.2019.27.253-258](https://doi.org/10.5455/aim.2019.27.253-258)] [Medline: [32055092](https://pubmed.ncbi.nlm.nih.gov/32055092/)]
40. Appari A, Johnson ME. Information security and privacy in healthcare: current state of research. *Int J Internet Enterp Manag* 2010;6(4):279-314. [doi: [10.1504/ijiem.2010.035624](https://doi.org/10.1504/ijiem.2010.035624)]
41. Yeng PK, Wolthusen SD, Yang B. Legal requirements towards enhancing the security of medical devices. *Int J Adv Comput Sci Appl* 2020;11(11):666-675. [doi: [10.14569/ijacsa.2020.0111181](https://doi.org/10.14569/ijacsa.2020.0111181)]
42. Donahue K, Rahman SM. Healthcare IT: is your information at risk? *Int J Netw Secur Appl* 2012 Sep 30;4(5):97-109 [FREE Full text] [doi: [10.5121/ijnsa.2012.4508](https://doi.org/10.5121/ijnsa.2012.4508)]
43. Coleman J. Assessing information security risk in healthcare organizations of different scale. *Int Congr Ser* 2004 Jun;1268:125-130. [doi: [10.1016/j.ics.2004.03.136](https://doi.org/10.1016/j.ics.2004.03.136)]
44. Vago S. Law and society. 10th edition. New York, NY, USA: Routledge; Aug 27, 2015.
45. Warren E. Legal, ethical, and professional issues in information security. Cengage Learning. 2018. URL: https://www.cengage.com/resource_uploads/downloads/1111138214_259148.pdf [accessed 2018-01-31]
46. Wellington KB. Cyberattacks on medical devices and hospital networks: legal gaps and regulatory solutions. *Santa Clara High Technol Law J* 2014;30:139.
47. Yang CM, Lin HC, Chang P, Jian WS. Taiwan's perspective on electronic medical records' security and privacy protection: lessons learned from HIPAA. *Comput Methods Programs Biomed* 2006 Jun;82(3):277-282. [doi: [10.1016/j.cmpb.2006.04.002](https://doi.org/10.1016/j.cmpb.2006.04.002)] [Medline: [16730852](https://pubmed.ncbi.nlm.nih.gov/16730852/)]
48. Section II: Management controls. Chapter 5: Computer Security Policy. National Institute of Standards and Technology. 2014. URL: <https://csrc.nist.gov/publications/nistpubs/800-12/800-12-html/chapter5.html> [accessed 2021-12-22]
49. Fites PE, Kratz MP. Information systems security: a practitioner's reference. New York, NY, USA: Van Nostrand Reinhold; 1994.
50. Lobel J. Foiling the system breakers: computer security and access control. New York, NY, USA: McGraw-Hill; 1986.
51. Kitchenham B, Pearl Brereton O, Budgen D, Turner M, Bailey J, Linkman S. Systematic literature reviews in software engineering – a systematic literature review. *Inf Softw Technol* 2009 Jan;51(1):7-15. [doi: [10.1016/j.infsof.2008.09.009](https://doi.org/10.1016/j.infsof.2008.09.009)]
52. Booth A, Sutton A, Papaioannou D. Systematic approaches to a successful literature review. London, UK: Sage Publications; 2016.
53. Khan RA, Khan SU. A preliminary structure of software security assurance model. In: Proceedings of the 13th International Conference on Global Software Engineering. 2018 Presented at: ICGSE '18; May 27-29, 2018; Gothenburg, Sweden p. 137-140. [doi: [10.1145/3196369.3196385](https://doi.org/10.1145/3196369.3196385)]
54. Petersen K, Vakkalanka S, Kuzniarz L. Guidelines for conducting systematic mapping studies in software engineering: an update. *Inf Softw Technol* 2015 Aug;64:1-18. [doi: [10.1016/j.infsof.2015.03.007](https://doi.org/10.1016/j.infsof.2015.03.007)]
55. Forskningsetikkens historie, akademisk frihet og publiseringsetikk. Forskningsetiske Komiteene. 2022. URL: <https://www.forskningsetikk.no/> [accessed 2022-03-22]
56. Kementerian Hukum dan HAM RI. 2021. URL: <https://peraturan.go.id> [accessed 2019-03-01]
57. Produk Hukor. Biro Hukum Dan Organisasi - Kementerian Kesehatan Republik Indonesia. 2021. URL: <http://hukor.kemkes.go.id/hukor/permenkes> [accessed 2022-03-22]
58. Ghana's Constitution of 1992 with Amendments through 1996. Constitute Project. 2021 Aug 26. URL: https://www.constituteproject.org/constitution/Ghana_1996.pdf [accessed 2022-03-22]
59. Policy documents. Ministry of health, Republic of Ghana. 2020. URL: <https://www.moh.gov.gh/policy-documents/> [accessed 2022-03-22]
60. Manuals Guidelines. Data protection Commission, Republic of Ghana. 2020. URL: <https://www.dataprotection.org.gh/> [accessed 2022-03-22]
61. Key Issues. General Data Protection Regulation. 2014. URL: <https://gdpr-info.eu/issues/> [accessed 2022-03-22]
62. Parsons K, McCormac A, Butavicius M, Pattinson M, Jerram C. Determining employee awareness using the Human Aspects of Information Security Questionnaire (HAIS-Q). *Comput Secur* 2014 May;42:165-176. [doi: [10.1016/j.cose.2013.12.003](https://doi.org/10.1016/j.cose.2013.12.003)]

63. H AIS-Q - a smart solution to cyber security. Department of Defence, Australian Government. 2017 Aug. URL: <https://www.dst.defence.gov.au/podcast/hais-q-smart-solution-cyber-security> [accessed 2022-03-22]
64. Code of conduct for information security and data protection in the healthcare and care services sector. Version 6.0. Norwegian Directorate of eHealth. 2020 Feb 5. URL: [https://www.ehelse.no/normen/documents-in-english/_attachment/download/f615c927-fdc5-4ac7-9c8e-99b0bd95ea9a:75611a1559201375f7cc62079fc6405ace48c841/Code%20of%20Conduct%20version%206.0%20\(PDF\).pdf](https://www.ehelse.no/normen/documents-in-english/_attachment/download/f615c927-fdc5-4ac7-9c8e-99b0bd95ea9a:75611a1559201375f7cc62079fc6405ace48c841/Code%20of%20Conduct%20version%206.0%20(PDF).pdf) [accessed 2021-06-22]
65. Mirkovic J, Bryhni H, Ruland CM. A framework for the development of ubiquitous patient support systems. In: Proceedings of the 6th International Conference on Pervasive Computing Technologies for Healthcare. 2012 Presented at: PervasiveHealth '12; May 21-24, 2012; San Diego, CA, USA p. 81-88. [doi: [10.4108/icst.pervasivehealth.2012.248594](https://doi.org/10.4108/icst.pervasivehealth.2012.248594)]
66. Act of 18 May 2001 N° 24 on Personal Health Data Filing Systems and the Processing of Personal Health Data (Personal Health Data Filing System Act). Ministry of Health and Care Services. 2014. URL: <https://tinyurl.com/2scfttpw> [accessed 2022-03-22]
67. Management system for information security. Helse. 2020. URL: <https://www.helse-sorost.no/informasjonsikkerhet-og-p> [accessed 2022-03-22]
68. Act of 2 July 1999 No. 64 relating to Health Personnel, etc. Ministry of Health and Care Services. 2002. URL: <https://www.regjeringen.no/no/dokumenter/act-of-2-july-1999-no-64-relating-to-hea/id107079/> [accessed 2022-03-22]
69. A framework for cybersecurity and cybercrime, and a contribution for peace, security and justice in cyberspace. In: Cyber Crime Law. Oslo, Norway: Juridika; May 18, 2017.
70. Herveg J. Data protection and the patient's right to safety. Eur J Health Law 2014 Jun;21(3):260-270. [doi: [10.1163/15718093-12341322](https://doi.org/10.1163/15718093-12341322)] [Medline: [25065033](https://pubmed.ncbi.nlm.nih.gov/25065033/)]
71. Aryanto K. Ensuring patient privacy in image data sharing for clinical research: design and implementation of rules and infrastructure. University of Gronigen. 2016. URL: https://pure.rug.nl/ws/portalfiles/portal/28139340/Chapter_1.pdf [accessed 2022-03-22]
72. European Union (EU). Directive (EU) 2016/1148 of the European Parliament and of the Council of 6 July 2016 concerning measures for a high common level of security of network and information systems across the Union. Official Journal of the European Union. 2016 Jul 19. URL: <https://eur-lex.europa.eu/legal-content/EN/TXT/PDF/?uri=OJ:L:2016:194:FULL&from=EN> [accessed 2022-03-22]
73. Quinn P, De Hert P. The Patients' Rights Directive (2011/24/EU) – Providing (some) rights to EU residents seeking healthcare in other Member States. Comput Law Secur Rev 2011 Sep;27(5):497-502. [doi: [10.1016/j.clsr.2011.07.010](https://doi.org/10.1016/j.clsr.2011.07.010)]
74. Kierkegaard P. Electronic health record: wiring Europe's healthcare. Comput Law Secur Rev 2011 Sep;27(5):503-515. [doi: [10.1016/j.clsr.2011.07.013](https://doi.org/10.1016/j.clsr.2011.07.013)]
75. Mossialos E, Thomson S, Ter Linden A. Information technology law and health systems in the European Union. Int J Technol Assess Health Care 2004;20(4):498-508. [doi: [10.1017/s0266462304001424](https://doi.org/10.1017/s0266462304001424)] [Medline: [15609802](https://pubmed.ncbi.nlm.nih.gov/15609802/)]
76. Hiller J, McMullen MS, Chumney WM, Baumer DL. Privacy and security in the implementation of health information technology (Electronic Health Records): U.S. and EU compared. BUJ Sci & Tech L 2011;17:1-40.
77. Pereira C, Oliveira C, Vilaca C, Ferreira A. Protection of clinical data: comparison of European with American legislation and respective technological applicability. In: Proceedings of the International Conference on Health Informatics. 2011 Presented at: HEALTHINF '11; January 26-29, 2011; Rome, Italy p. 567-570. [doi: [10.5220/0003165505670570](https://doi.org/10.5220/0003165505670570)]
78. Di Iorio CT, Carinci F, Oderkirk J, Smith D, Siano M, de Marco DA, et al. Assessing data protection and governance in health information systems: a novel methodology of Privacy and Ethics Impact and Performance Assessment (PEIPA). J Med Ethics 2020 Mar 27:105948. [doi: [10.1136/medethics-2019-105948](https://doi.org/10.1136/medethics-2019-105948)] [Medline: [32220868](https://pubmed.ncbi.nlm.nih.gov/32220868/)]
79. Todde M, Beltrame M, Marceglia S, Spagno C. Methodology and workflow to perform the Data Protection Impact Assessment in healthcare information systems. Inform Med Unlocked 2020;19:100361. [doi: [10.1016/j.imu.2020.100361](https://doi.org/10.1016/j.imu.2020.100361)]
80. Sousa M, Ferreira DN, Santos-Pereira C, Bacelar G, Frade S, Pestana O, et al. openEHR based systems and the General Data Protection Regulation (GDPR). Stud Health Technol Inform 2018;247:91-95. [Medline: [29677929](https://pubmed.ncbi.nlm.nih.gov/29677929/)]
81. Paluzzi M. Paying prices for swiped devices: addressing the issue of medical identity theft from unencrypted stolen laptops. University of Illinois Law Review. 2019 Sep 1. URL: <https://www.illinoislawreview.org/wp-content/uploads/2019/09/Paluzzi.pdf> [accessed 2022-04-12]
82. Yuan B, Li J. The policy effect of the General Data Protection Regulation (GDPR) on the digital public health sector in the European Union: an empirical investigation. Int J Environ Res Public Health 2019 Mar 25;16(6):1070 [FREE Full text] [doi: [10.3390/ijerph16061070](https://doi.org/10.3390/ijerph16061070)] [Medline: [30934648](https://pubmed.ncbi.nlm.nih.gov/30934648/)]
83. Pinem AA, Fajrina HR, Shandyaduhita PI, Handayani PW, Hidayanto AN. Barriers for integration between hospitals and the Ministry of Health in Indonesia. In: Managing Intellectual Capital and Innovation for Sustainable and Inclusive Society: Proceedings of the MakeLearn and TIIM Joint International Conference. 2015 Presented at: MakeLearn and TIIM '16; May 27-29, 2015; Bari, Italy p. 807-817.
84. Anderson RJ. A security policy model for clinical information systems. In: Proceedings 1996 IEEE Symposium on Security and Privacy. 1996 Presented at: SECPRI '96; May 6-8, 1996; Oakland, CA, USA p. 30-43. [doi: [10.1109/secpri.1996.502667](https://doi.org/10.1109/secpri.1996.502667)]

85. Mennerat F, Chabriaux J. Do EHR communication standards account for imaging communication needs? In: Lemke HU, Inamura K, Doi K, Vannier MW, Farman AG, Reiber JH, editors. CARS 2002 Computer Assisted Radiology and Surgery. Berlin, Heidelberg: Springer; 2002:647-650.
86. Hu Y, Afzal J. Achieving eHealth interoperability via peer-to-peer communication using JXTA technology. Blekinge Institute of Technology. 2010 May. URL: <https://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.471.702&rep=rep1&type=pdf> [accessed 2022-03-22]
87. Sari PK, Handayani PW, Hidayanto AN. Security value issues on eHealth implementation in Indonesia. IOP Conf Ser Mater Sci Eng 2020 Jul 01;879(1):012040. [doi: [10.1088/1757-899x/879/1/012040](https://doi.org/10.1088/1757-899x/879/1/012040)]
88. Seddon JJ, Currie WL. Cloud computing and trans-border health data: Unpacking U.S. and EU healthcare regulation and compliance. Health Policy Technol 2013 Dec;2(4):229-241. [doi: [10.1016/j.hlpt.2013.09.003](https://doi.org/10.1016/j.hlpt.2013.09.003)]
89. Bincoletto G. Data protection issues in cross-border interoperability of Electronic Health Record systems within the European Union. Data Policy 2020 Mar 30;2:e3. [doi: [10.1017/dap.2020.2](https://doi.org/10.1017/dap.2020.2)]
90. Rath TA, Colin JN. Adaptive risk-aware access control model for internet of things. In: Proceedings of the 2017 International Workshop on Secure Internet of Things. 2017 Presented at: SIoT '17; September 15, 2017; Oslo, Norway p. 40-49. [doi: [10.1109/siot.2017.00010](https://doi.org/10.1109/siot.2017.00010)]
91. Pulkkis G, Karlsson J, Westerlund M, Tana J. Secure and reliable Internet of Things systems for healthcare. In: Proceedings of the IEEE 5th International Conference on Future Internet of Things and Cloud. 2017 Presented at: FiCloud '17; Aug 21-23, 2017; Prague, Czech Republic p. 169-176. [doi: [10.1109/ficloud.2017.50](https://doi.org/10.1109/ficloud.2017.50)]
92. Natsiavas P, Rasmussen J, Voss-Knude M, Votis K, Coppolino L, Campegiani P, et al. Comprehensive user requirements engineering methodology for secure and interoperable health data exchange. BMC Med Inform Decis Mak 2018 Oct 16;18(1):85 [FREE Full text] [doi: [10.1186/s12911-018-0664-0](https://doi.org/10.1186/s12911-018-0664-0)] [Medline: [30326890](https://pubmed.ncbi.nlm.nih.gov/30326890/)]
93. Saranto K, Kivekäs E, Kinnunen UM, Palojoki S. Lack of patient data privacy challenges patient safety. Stud Health Technol Inform 2018;251:163-166. [Medline: [29968628](https://pubmed.ncbi.nlm.nih.gov/29968628/)]
94. Andersen MR, Storm HH, Eurocourse Work Package 2 Group. Cancer registration, public health and the reform of the European data protection framework: abandoning or improving European public health research? Eur J Cancer 2015 Jun;51(9):1028-1038. [doi: [10.1016/j.ejca.2013.09.005](https://doi.org/10.1016/j.ejca.2013.09.005)] [Medline: [24120502](https://pubmed.ncbi.nlm.nih.gov/24120502/)]
95. Callens S, Cierkens K. Legal aspects of E-HEALTH. Stud Health Technol Inform 2008;141:47-56. [Medline: [18953124](https://pubmed.ncbi.nlm.nih.gov/18953124/)]
96. Nordland O, Sujana MA, Koornneef F, Bernsmed K. Assurance requirements for networked medical sensor applications. In: Proceedings of the 2015 IEEE International Conference on Industrial Technology. 2015 Presented at: ICIT '15; March 17-19, 2015; Seville, Spain p. 1838-1844. [doi: [10.1109/icit.2015.7125364](https://doi.org/10.1109/icit.2015.7125364)]
97. Kierkegaard P. E-Prescription across Europe. Health Technol 2012 Dec 20;3(3):205-219. [doi: [10.1007/s12553-012-0037-0](https://doi.org/10.1007/s12553-012-0037-0)]
98. Klein GO. Standardization strategy from a European perspective. Int J Med Inform 1998 Feb;48(1-3):67-70. [doi: [10.1016/s1386-5056\(97\)00112-3](https://doi.org/10.1016/s1386-5056(97)00112-3)] [Medline: [9600406](https://pubmed.ncbi.nlm.nih.gov/9600406/)]
99. Christiansen EK, Skipenes E, Hausken MF, Skeie S, Østbye T, Iversen MM. Shared electronic health record systems: key legal and security challenges. J Diabetes Sci Technol 2017 Nov;11(6):1234-1239 [FREE Full text] [doi: [10.1177/1932296817709797](https://doi.org/10.1177/1932296817709797)] [Medline: [28560899](https://pubmed.ncbi.nlm.nih.gov/28560899/)]
100. Brox EA. Information security in distributed health information systems in Scandinavia: a comparative study of external conditions and solutions for exchange and sharing of sensitive health information in Denmark, Norway and Sweden. Norwegian University of Science and Technology. 2006 Jun. URL: <https://ntnuopen.ntnu.no/ntnu-xmlui/handle/11250/251121> [accessed 2022-03-22]
101. Kalra D. Electronic Health Record Standards. Yearb Med Inform 2018 Mar 07;15(01):136-144. [doi: [10.1055/s-0038-1638463](https://doi.org/10.1055/s-0038-1638463)]
102. Nalin M, Baroni I, Faiella G, Romano M, Matrisciano F, Gelenbe E, et al. The European cross-border health data exchange roadmap: case study in the Italian setting. J Biomed Inform 2019 Jun;94:103183 [FREE Full text] [doi: [10.1016/j.jbi.2019.103183](https://doi.org/10.1016/j.jbi.2019.103183)] [Medline: [31009760](https://pubmed.ncbi.nlm.nih.gov/31009760/)]
103. Healy L, Lubeck P. Patient information privacy and security in Ghana: current policy and suggestions for the future. University of California Santa Cruz. 2012. URL: <https://dca.ucsc.edu/dca/winners/2012/159> [accessed 2022-04-12]
104. Setianto YD, Wahyuningrum SE. Medical device authentication and authorization protocol in Indonesian telemedicine systems. In: Proceedings of the 4th International Conference on Information Technology. 2019 Presented at: InCIT '10; October 24-25, 2019; Bangkok, Thailand p. 89-93. [doi: [10.1109/incit.2019.8912058](https://doi.org/10.1109/incit.2019.8912058)]
105. Santoso I, Siahaan IS, Suharjito. Privacy modelling of sensitive data in universal healthcare coverage in Indonesia. In: Proceedings of the 11th International Conference on Knowledge, Information and Creativity Support Systems. 2016 Presented at: KICSS '16; November 10-12, 2016; Yogyakarta, Indonesia p. 1-6. [doi: [10.1109/KICSS.2016.7951437](https://doi.org/10.1109/KICSS.2016.7951437)]
106. Høstgaard AM, Bertelsen P, Nøhr C. Methods to identify, study and understand end-user participation in HIT development. BMC Med Inform Decis Mak 2011 Sep 28;11:57 [FREE Full text] [doi: [10.1186/1472-6947-11-57](https://doi.org/10.1186/1472-6947-11-57)] [Medline: [21955493](https://pubmed.ncbi.nlm.nih.gov/21955493/)]
107. Samonas S, Coss D. The CIA strikes back: redefining confidentiality, integrity and availability in security. J Inf Syst Secur 2014;10(3):21-45.
108. What are the GDPR consent requirements? General Data Protection Regulation. 2019. URL: <https://gdpr.eu/gdpr-consent-requirements/> [accessed 2022-03-22]

109. Act of 18 May 2001 No. 24 on Personal Health Data Filing Systems and the Processing of Personal Health Data (Personal Health Data Filing System Act). Ministry of Justice and Public Security. 2017. URL: https://www.datatilsynet.no/globalassets/global/english/personal_health_data_filing_system_act_20100907.pdf [accessed 2022-03-22]
110. Yarney L, Buabeng T, Baidoo D, Bawole JN. Operationalization of the Ghanaian patients' charter in a peri-urban public hospital: voices of healthcare workers and patients. *Int J Health Policy Manag* 2016 Sep 01;5(9):525-533 [[FREE Full text](#)] [doi: [10.15171/ijhpm.2016.42](https://doi.org/10.15171/ijhpm.2016.42)] [Medline: [27694679](https://pubmed.ncbi.nlm.nih.gov/27694679/)]
111. ACT 2008-06-20 no. 44: Act on medical and health research (the Health Research Act). Universitetet I Oslo. 2008. URL: <https://app.uio.no/ub/ujur/oversatte-lover/data/lov-20080620-044-eng.pdf> [accessed 2022-03-22]
112. Yousefi A. The impact of information and communication technology on economic growth: evidence from developed and developing countries. *Econ Innov New Technol* 2011 Sep;20(6):581-596. [doi: [10.1080/10438599.2010.544470](https://doi.org/10.1080/10438599.2010.544470)]
113. Ngwa W, Olver I, Schmeler KM. The use of health-related technology to reduce the gap between developed and undeveloped regions around the globe. *Am Soc Clin Oncol* 2020 May(40):227-236. [doi: [10.1200/edbk_288613](https://doi.org/10.1200/edbk_288613)]
114. Miltgen CL, Peyrat-Guillard D. Cultural and generational influences on privacy concerns: a qualitative study in seven European countries. *Eur J Inf Syst* 2019 Jan 28;23(2):103-125. [doi: [10.1057/ejis.2013.17](https://doi.org/10.1057/ejis.2013.17)]
115. Yeng PK, Yang B, Solvoll T, Nimbe P, Weyori BA. Web Vulnerability Measures for SMEs. Norsk Informasjonssikkerhetskonferanse 2019:1-16 [[FREE Full text](#)]

Abbreviations

CEO: Chief Executive Officer

CIA: confidentiality, integrity, and availability

EHR: electronic health record

EISP: enterprise or organizational information security policy

EU: European Union

GDPR: General Data Protection Regulation

ICT: information and communication technology

ISO: International Organization for Standardization

ISSP: issue-specific security policy

IT: information technology

PRISMA: Preferred Reporting Items for Systematic Reviews and Meta-Analyses

Edited by A Kushniruk; submitted 29.04.21; peer-reviewed by G Klein, F Ghezelbash, R Ciorap; comments to author 30.07.21; revised version received 25.09.21; accepted 03.02.22; published 25.05.22.

Please cite as:

Yeng PK, Fauzi MA, Sun L, Yang B

Assessing the Legal Aspects of Information Security Requirements for Health Care in 3 Countries: Scoping Review and Framework Development

JMIR Hum Factors 2022;9(2):e30050

URL: <https://humanfactors.jmir.org/2022/2/e30050>

doi: [10.2196/30050](https://doi.org/10.2196/30050)

PMID: [35612891](https://pubmed.ncbi.nlm.nih.gov/35612891/)

©Prosper Kandabongee Yeng, Muhammad Ali Fauzi, Luyi Sun, Bian Yang. Originally published in JMIR Human Factors (<https://humanfactors.jmir.org>), 25.05.2022. This is an open-access article distributed under the terms of the Creative Commons Attribution License (<https://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Human Factors, is properly cited. The complete bibliographic information, a link to the original publication on <https://humanfactors.jmir.org>, as well as this copyright and license information must be included.

Review

Digital Health Opportunities to Improve Primary Health Care in the Context of COVID-19: Scoping Review

Cícera Renata Diniz Vieira Silva¹, PhD; Rayssa Horácio Lopes², MSc; Osvaldo de Goes Bay Jr³, PhD; Claudia Santos Martiniano⁴, PhD; Miguel Fuentealba-Torres⁵, PhD; Ricardo Alexandre Arcêncio⁶, PhD; Luís Velez Lapão^{7,8}, PhD; Sonia Dias⁹, PhD; Severina Alice da Costa Uchoa², PhD

¹Faculty of Health Sciences, Federal University of Rio Grande do Norte, Natal, Brazil

²Department of Collective Health, Federal University of Rio Grande do Norte, Natal, Brazil

³Faculty of Health Sciences, Trairi, Federal University of Rio Grande do Norte, Santa Cruz, Brazil

⁴Department of Nursing, State University of Paraíba, Campina Grande, Brazil

⁵Universidad de los Andes, Santiago, Chile

⁶Department of Maternal Infant Nursing and Public Health, University of São Paulo, Ribeirão Preto, Brazil

⁷Instituto de Higiene e Medicina Tropical, Comprehensive Health Research Center, Universidade Nova de Lisboa, Lisbon, Portugal

⁸Unidade de Investigação e Desenvolvimento em Engenharia Mecânica e Industrial, Universidade Nova de Lisboa, Lisbon, Portugal

⁹Escola Nacional de Saúde Pública, Comprehensive Health Research Center, Universidade Nova de Lisboa, Lisbon, Portugal

Corresponding Author:

Cícera Renata Diniz Vieira Silva, PhD

Faculty of Health Sciences

Federal University of Rio Grande do Norte

601 General Gustavo Cordeiro de Faria Street

Natal, 59012-570

Brazil

Phone: 55 84 3221 0862

Email: renatadiniz_enf@yahoo.com.br

Abstract

Background: The COVID-19 pandemic brought social, economic, and health impacts, requiring fast adaptation of health systems. Although information and communication technologies were essential for achieving this objective, the extent to which health systems incorporated this technology is unknown.

Objective: The aim of this study was to map the use of digital health strategies in primary health care worldwide and their impact on quality of care during the COVID-19 pandemic.

Methods: We performed a scoping review based on the Joanna Briggs Institute manual and guided by the PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-analyses) Extension for Scoping Reviews. A systematic and comprehensive three-step search was performed in June and July 2021 in multidisciplinary health science databases and the gray literature. Data extraction and eligibility were performed by two authors independently and interpreted using thematic analysis.

Results: A total of 44 studies were included and six thematic groups were identified: characterization and geographic distribution of studies; nomenclatures of digital strategies adopted; types of information and communication technologies; characteristics of digital strategies in primary health care; impacts on quality of care; and benefits, limitations, and challenges of digital strategies in primary health care. The impacts on organization of quality of care were investigated by the majority of studies, demonstrating the strengthening of (1) continuity of care; (2) economic, social, geographical, time, and cultural accessibility; (3) coordination of care; (4) access; (5) integrality of care; (6) optimization of appointment time; (7) and efficiency. Negative impacts were also observed in the same dimensions, such as reduced access to services and increased inequity and unequal use of services offered, digital exclusion of part of the population, lack of planning for defining the role of professionals, disarticulation of actions with real needs of the population, fragile articulation between remote and face-to-face modalities, and unpreparedness of professionals to meet demands using digital technologies.

Conclusions: The results showed the positive and negative impacts of remote strategies on quality of care in primary care and the inability to take advantage of the potential of technologies. This may demonstrate differences in the organization of fast and urgent implementation of digital strategies in primary health care worldwide. Primary health care must strengthen its response

capacity, expand the use of information and communication technologies, and manage challenges using scientific evidence since digital health is important and must be integrated into public service.

(*JMIR Hum Factors* 2022;9(2):e35380) doi:[10.2196/35380](https://doi.org/10.2196/35380)

KEYWORDS

digital health; telehealth; telemedicine; primary health care; quality of care; COVID-19; pandemic; science database; gray literature

Introduction

Quality in health care is a multidimensional concept related to how offered services increase the probability of desired health outcomes. Health care quality also permeates correct care at the right time and in a coordinated manner, responding to the needs and preferences of service users, and reducing damage and wasted resources through a continuous and dynamic process [1]. Quality of care approximates health services to the population and has three dimensions: technical (accuracy of actions and the way they are performed), interpersonal (social and psychological relationships between care providers and users), and organizational (conditions in which services are offered, including globalization and continuity of care, coverage, coordination of actions, access, and accessibility to services) [2-4].

The COVID-19 pandemic led to immediate and profound social, economic, and health impacts, and required fast adaptation of health systems focusing on quality. Health systems, particularly primary health care (PHC), were pushed to maintain care routines, which required changes to maintain access and continuous management of health problems. This was possible owing to the creativity and innovation of professionals and managers, who introduced or expanded the use of information and communication technology (ICT) in the critical initial phase of the pandemic, where lack of coordination has negatively influenced access to health care [5].

ICT use has digital health as a great exponent in remote care strategies. This term is historically addressed as telemedicine or telehealth, which refers to communication and interaction tools between health care professionals and patients that provide remote health services and care as alternative to face-to-face appointments [6-8].

The use of telephones to answer doubts of patients, videos or text messages through mobile apps, and social media are helpful strategies for expanding the scope of health care by enabling population access. ICT also reduces the distance between patients and health professionals (eg, rural areas lacking health professionals), and facilitates appointment scheduling and renewal of prescriptions, thereby changing the professional-patient relationship and expanding personal health management [6,7,9-11].

The COVID-19 pandemic became a catalyst for expanding ICT worldwide [12]. Although digital health was recommended by the World Health Organization (WHO) [13-15] to reduce geographic barriers, its use increased only during the pandemic to maintain or increase access to health care, fight the pandemic, minimize economic impacts, and enable continuity of remote care [16,17].

Technological evolution may accelerate health care and improve access in the context of public health preparedness and response to outbreaks and emergencies. Despite these advances, the pandemic was challenging for health systems, mainly due to the lack of integration of technologies [17,18]. Considering the relevance of the topic for health and the wide use of ICT in PHC during the pandemic, we sought to gather knowledge about the quality of PHC using digital technologies. Therefore, the aim of this study was to map the use of digital health strategies in PHC worldwide and their impact on quality of care in the context of the COVID-19 pandemic.

Methods

Design

This scoping review was performed based on the Joanna Briggs Institute (JBI) manual [19] and guided by PRISMA-ScR (Preferred Reporting Items for Systematic Reviews and Meta-analyses Extension for Scoping Reviews) guidelines [20]. We also followed the steps proposed by Arksey and O'Malley [21] and Levac et al [22]: formulation of research questions, identification of relevant studies, study selection, data extraction and coding, analysis and interpretation of results, and consultation with stakeholders.

Ethics Approval

The study was approved by the research ethics committee of the Faculty of Health Sciences of Trairí, Federal University of Rio Grande do Norte (CAAE: 47473121.3.0000.5568), and direct participation of participants involved in the study occurred only during consultation with stakeholders. The methodology used was previously reported in a protocol [23]. The term "telemedicine" used in the protocol [23] was replaced by "digital health" in this scoping review since it was considered to be more appropriate to reflect the broad scope of the study.

Formulation of Research Questions

Study questions were defined by consensus among the authors and were formulated using the PCC (Population, Concept, and Context) mnemonic and the respective results of interest [19]: (1) Which countries used digital health in PHC in response to the COVID-19 pandemic? (2) What options of ICT were used in PHC in the context of the COVID-19 pandemic? (3) What is the impact of digital health on quality of health care delivery in PHC in the context of the COVID-19 pandemic?

Identification of Relevant Studies

The following multidisciplinary health science databases were searched for relevant articles: MEDLINE/PubMed, Scopus, Web of Science, CINAHL, Embase, and LILACS. For gray literature, we consulted Google Scholar, WHO Global Research

on Coronavirus Disease, PAHO Technical Documents and Research Evidence on COVID-19, Cochrane Library, medRxiv, SciELO Preprints, preprints.org, Open Grey, and Grey Literature Report.

The following types of studies and documents that addressed the research questions, focused on the use of remote strategies in PHC during the COVID-19 pandemic, and were available in full text were included: primary studies with quantitative, qualitative, or a mixed approach; experience reports; case reports; intervention studies; preprints; guidelines; manuals; reports; and government documents. No date or language filters were applied. Duplicate studies, protocols, literature reviews, opinion letters, and editorials were excluded.

Study Selection

The search was performed between June 14 and July 14, 2021, using a three-step search strategy [24]: (1) exploratory search in two databases to identify descriptors and keywords, followed by construction of the search strategy, which was improved by a librarian using the Extraction, Conversion, Combination, Construction, and Use model [25]; (2) definition and search in all databases; and (3) manual search for additional sources in references of selected studies. The detailed search strategies are presented in [Multimedia Appendix 1](#).

Study selection followed the PRISMA steps [26]: identification, screening, eligibility, and inclusion. A pilot study was independently conducted by two authors (CRDVS and RHL) using Rayyan software [27] to verify blinding, exclusion of duplicates, and selection of studies by titles and abstracts. Subsequently, full texts and reference lists of included studies were analyzed. For studies that did not meet inclusion criteria, a third author (SACU) was consulted.

Data Extraction and Coding

Data extraction and coding ensured the consistency and reliability of results. Two authors (CRDVS and RHL) independently extracted all relevant data using an extraction form based on the JBI template [24], which was adapted by the

authors, containing the following information: characterization of studies (first author, year, journal, country, type of study, participants); names of digital strategies adopted; types of ICT; characteristics of digital strategies in PHC; impacts on quality of care; and benefits, limitations, and challenges of digital strategies in PHC.

The database was organized in a Microsoft Excel 2016 spreadsheet and is provided for consultation in [Multimedia Appendix 2](#).

Analysis and Interpretation of Results

Data were analyzed qualitatively (narrative analysis) and quantitatively (absolute and relative frequencies). Thematic analysis [28] was structured based on familiarization with data, generation of initial codes, search for topics, review of topics, definition and naming of topics, and implications of studies. Results and narrative analyses are reported in tables and figures.

Consultation With Stakeholders

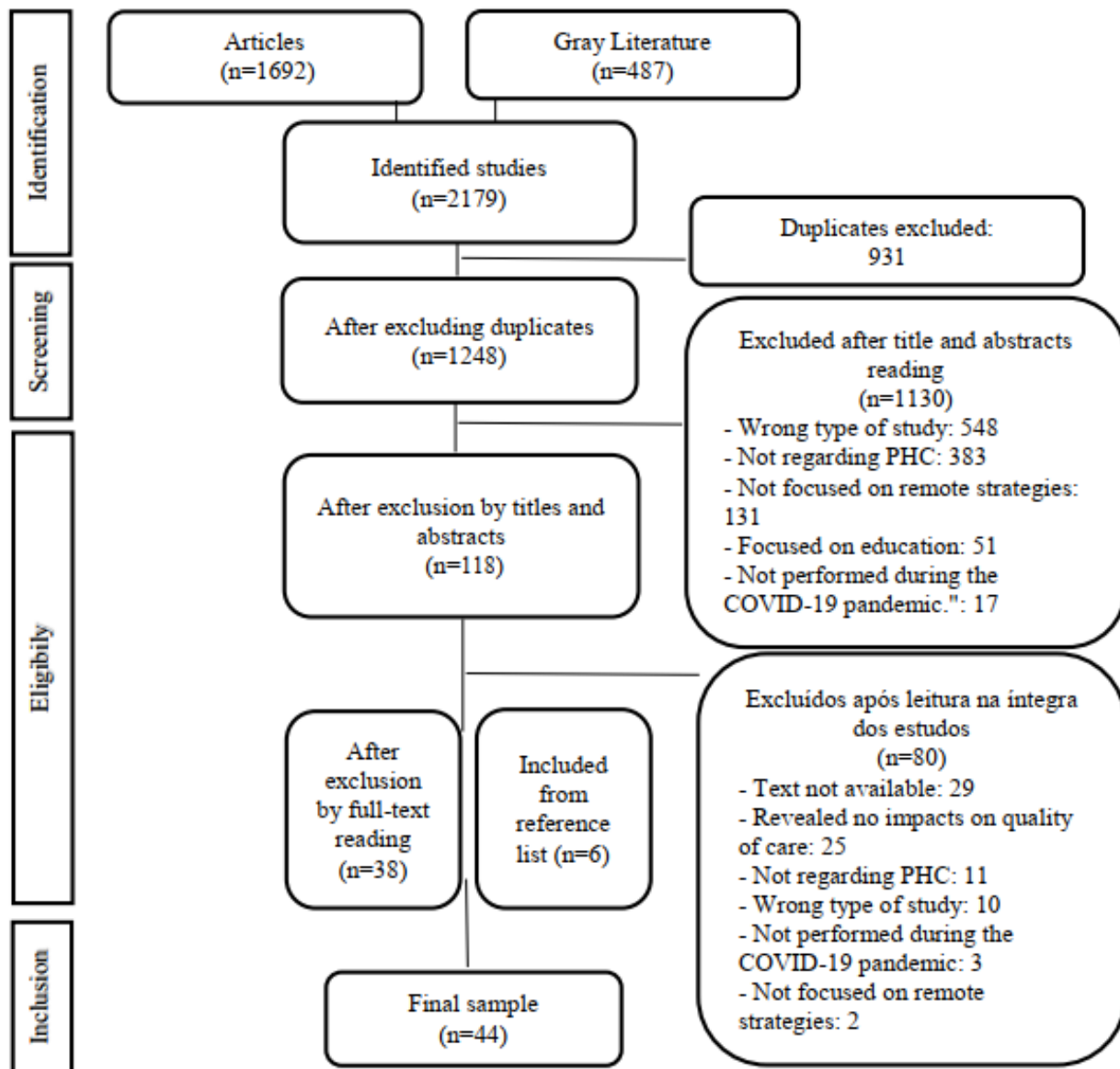
Results of this review were presented to five stakeholders (ie, researchers with experience in digital health, ICT in health, and PHC) to fulfill the following objectives recommended by Levac et al [22]: preliminary sharing of study findings, considered a mechanism for transferring and exchanging knowledge, and development of effective dissemination strategies and ideas for future studies. The form questions are provided in [Multimedia Appendix 3](#).

Results

Included Studies

A total of 2179 publications were identified (1692 peer-reviewed articles and 487 gray literature documents). After excluding duplicates, analysis of titles and abstracts, and full-text reading, 38 studies were included. The manual search of reference lists added 6 studies, resulting in a total of 44 publications for analysis ([Figure 1](#)). All included studies demonstrated the impacts of remote strategies on quality of care in PHC in the context of COVID-19.

Figure 1. Flowchart of study selection for scoping review adapted from the Preferred Reporting Items for Systematic Review and Meta-Analyses (PRISMA).



Characterization and Geographic Distribution of Studies

The studies included were mostly published during 2021 (28/44, 64%). Among the 44 articles, 27 (61%) used a cross-sectional design, 6 (14%) used qualitative investigation, 6 (14%) used mixed methods, 2 (5%) were cohort studies, 1 (2%) was an experience report, 1 (2%) was a case report, and 1 (2%) was an intervention study. The sample consisted mainly of patients (n=19, 43%), health professionals (n=13, 30%), medical or consultation records (n=9, 21%), and documented interviews with patients or health professionals (n=3, 7%).

The studies covered 18 countries that used digital strategies in PHC; 18 studies were performed in North America (United States [29-43] and Canada [44-46]), 4 studies were performed in South America (Brazil [47-50]), 14 studies were performed in Europe (England [51-53], United Kingdom [54,55], Spain [56,57], Belgium [58,59], Norway [60], Portugal [61], Romania [62], Germany [63], and Poland [64]), four studies were performed in Asia (Israel [65], Oman [66], Saudi Arabia [67], and Iran [68]), and four studies were performed in Oceania (Australia [69-71] and New Zealand [72]). The characteristics of studies and distribution of countries that used digital strategies in PHC are described in Table 1 and Figure 2, respectively.

Table 1. Characteristics of the included studies.

Reference	Source	Country	Study design	Participants/sample
Alexander et al [29]	JAMA Network Open	United States	Cross-sectional	National audit of consultations (n=117.9 million)
Schweiberger et al [30]	Journal of Medical Internet Research	United States	Cross-sectional	Electronic medical records (n=45) and physicians (n=121)
Olayiwola et al [31]	JMIR Public Health Surveillance	United States	Cross-sectional	Consultation records (n=3617)
Atherly et al [32]	JMIR Public Health Surveillance	United States	Cross-sectional	Patients (n=1694)
Judson et al [33]	Journal of the American Medical Informatics Association	United States	Cross-sectional	Consultation records (n=1129)
Mills et al [34]	Journal of the American Health Association	United States	Cross-sectional	Patients (n=587)
Tarn et al [35]	Journal of the American Board of Family Medicine	United States	Cross-sectional	Medical records (n=202)
Adepoju et al [36]	Journal of Health Care for the Poor and Underserved	United States	Cross-sectional	Health workers (n=1344)
Ritchie et al [37]	Journal of the American Medical Directors Association	United States	Mixed methods	Health workers (n=79)
Drerup et al [38]	Telemedicine Journal and e-Health	United States	Cross-sectional	Patients (n=65)
Kalicki et al [39]	Journal of the American Geriatrics Society	United States	Cross-sectional	Medical records (n=873)
Chang et al [40]	Milbank Quarterly	United States	Cross-sectional	Health workers (n=918)
Thies et al [41]	Journal of Primary Care & Community Health	United States	Cross-sectional	Health workers (n=655)
Godfrey et al [42]	Contraception	United States	Cross-sectional	Medical records (n=534)
Juarez-Reyes et al [43]	Therapeutic Advances in Chronic Disease	United States	Qualitative investigation	Patients (n=6)
Bui et al [44]	Hamilton Family Health Team	Canada	Cross-sectional	Clinicians (n=126) and nurses (n=6)
Mohammed et al [45]	PLoS One	Canada	Cross-sectional	Clinicians (n=163) and nurses (n=37)
Donnelly et al [46]	BMC Family Practice	Canada	Mixed methods	Health workers (n=473)
Castro et al [47]	Revista Brasileira de Medicina da Família e da Comunidade	Brazil	Cross-sectional	Consultation records (n=329)
Dimer et al [48]	CoDAS	Brazil	Experience report	Consultation records (n=17)
Queiroz et al [49]	Acta Diabetologica	Brazil	Cohort	Patients (n=627)
Silva et al [50]	Ciência e Saúde Coletiva	Brazil	Cross-sectional	Clinicians and nurses (n=7054)
Sahni et al [51]	Cureus	England	Cross-sectional	Clinicians (n=312)
Leung et al [52]	BMJ Open Quality	England	Intervention study	Patients (n=12)
Tuijt et al [53]	British Journal of General Practice	England	Qualitative investigation	Patients (n=30) and caregivers (n=31)
Salisbury et al [54]	Journal of Medical Internet Research	United Kingdom	Mixed methods	Patients (n=1452) and health workers (n=12)
Murphy et al [55]	British Journal of General Practice	United Kingdom	Mixed methods	Medical records (n=350,966) and health workers (n=87)
Llamosas et al [56]	Physiotherapy	Spain	Case report	Patient (n=1)
Coronado-Vázquez et al [57]	Journal of Personalized Medicine	Spain	Cohort	Patients (n=166)
Morreel et al [58]	PLoS One	Belgium	Cross-sectional	Home visit records (n=15,655)
Verhoeven et al [59]	BMJ Open	Belgium	Qualitative investigation	Patients (n=132)

Reference	Source	Country	Study design	Participants/sample
Johnsen et al [60]	Journal of Medical Internet Research	Norway	Cross-sectional	Clinicians (n=1237)
Lapão et al [61]	Journal of Medical Internet Research	Portugal	Mixed methods	Patients (n=35)
Florea et al [62]	International Journal of General Medicine	Romania	Cross-sectional	Clinicians (n=108)
Mueller et al [63]	JMIR Medical Informatics	Germany	Qualitative investigation	Patients (n=20)
Kludacz-Alessandri et al [64]	PLoS One	Poland	Cross-sectional	Patients (n=100)
Zeltzer et al [65]	National Bureau of Economic Research (NBER)/NBER Working Paper Series	Israel	Cross-sectional	Records from clinicians (n=4293) and patients (n=3.7 million)
Hasani et al [66]	Journal of Primary Care & Community Health	Oman	Qualitative investigation	Clinicians (n=22)
Alharbi et al [67]	Journal of Family and Community Medicine	Saudi Arabia	Cross-sectional	Patients (n=439)
Jannati et al [68]	International Journal of Medical Informatics	Iran	Cross-sectional	Patients (n=400)
Isautier et al [69]	Journal of Medical Internet Research	Australia	Cross-sectional	Patients (n=596)
Javanparast et al [70]	BMC Family Practice	Australia	Qualitative investigation	Patients (n=30)
Ervin et al [71]	Australian Journal of Primary Health	Australia	Cross-sectional	Clinicians (n=24)
Imlach et al [72]	BMC Family Practice	New Zealand	Mixed methods	Patients (n=1010)

Figure 2. Distribution of countries that used digital strategies in primary health care. Numbers represent the number of studies performed in each country.



Nomenclatures of Adopted Digital Strategies

Nomenclatures regarding remote care strategies varied considerably among studies, with the terms “telehealth” [30,33,36,44,45,47,54,55,60,63,64,67-69,72] and “telemedicine” [29,31,32,38,39,46,47,51,59,61,63,69,71] being the most frequent. The following terms were also mentioned:

teleconsultation [40,58,71], virtual visit [41,48,58], virtual health/eHealth [35,51], remote consultation [37,50,56,65], electronic consultation [35,62], telephone follow-up [35,66], video visit [35,70], video consultation [34,49], online consultation [69], virtual care [53], web-based video consultation [69], digital monitoring [72], nonpresential consultation [52],

and remote self-monitoring [43]. Figure 3 shows the word cloud representing the most commonly used nomenclatures.

Figure 3. Word cloud with nomenclatures used to refer to digital strategies in primary health care.



Types of ICT Employed

A total of 39 of the 44 studies (89%) mentioned the types of ICT used in PHC. Telephone calls had the highest number of records (29/39, 74%) [30,31,33-38,40,44-48, 50,52,53,55-60,62,65,66,69,70,72], followed by video calls (25/39, 64%) [30,31,33-41,43-48,55,60,62-65,69,72], patient portal (11/39, 28%) [31,33,35-37,40,42,44,58,61,72], smartphone apps (5/39, 13%) [44,49,52,54,68], text messages (3/39, 8%) [35,46,60], email (3/39, 8%) [46,62,72], electronic medical record (2/39, 5%) [31,72], and social networks (1/39, 3%) [46]. We highlight that many studies used more than one type of technology, mainly phone and video calls.

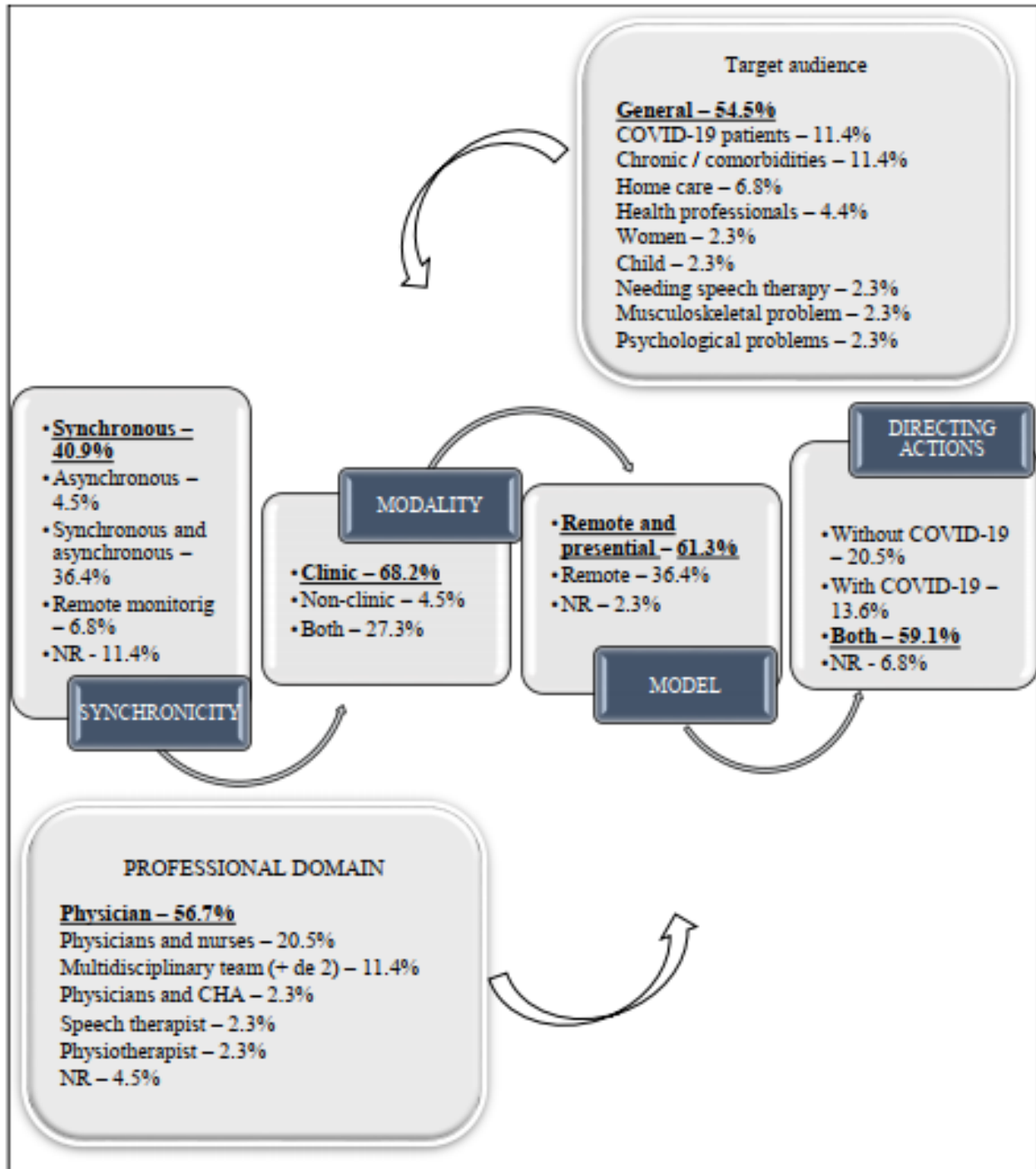
Moreover, the following electronic platforms and apps were used to conduct services: WhatsApp, Updox, Epic MyChart, Doximity, Facetime, Skype, Zoom, Telegram, iCARE-DATA, Babylon GP at Hand (BGPaH), EyerCloud, DRiQ, Aid Access, Telus PS Suite, eVisit Ontario Telemedicine Network, and Multimorbidity Management Health Information System (METHIS).

Characteristics of Digital Strategies in PHC

We analyzed the target audience, professionals involved, direction, synchronicity, and modality and model of actions in PHC (Figure 4).

Actions were mostly directed to the general public (ie, any health status or characteristic) [29,31,32,34-36,38,40-42, 44,46,48,49,51,53,55,56,58-60,63,67,71]. Regarding professionals who conducted the actions, the majority were clinicians [29-31, 34, 36-38, 42, 44-46, 49-51, 54, 55, 57, 58, 60, 63-66, 69, 71, 72] and nurses [35,39,41,43,47,48,56,70,72]. Actions were directed toward people with and without COVID-19 [31,32,34,36,38,40,42, 44,49,51,53-56,58-63, 65,67,68,70-72]. Synchronous interaction [31-33, 40,42,44-46,51,52,54,63-66,68,70,71] was the most frequently reported interaction type. The clinical modality was the most commonly reported [29-31, 33, 36-38, 40-46, 48-52, 54, 58-60, 63, 65-67, 69, 71, 72], referring to the following actions: consultations, renewal of medical prescriptions, exams, follow-up, health guidelines, issuance of certificates, treatments, screening, monitoring, diagnosis, management of chronic conditions, referrals, clinical self-monitoring, and risk classification. Remote consultation associated with in-person actions was the most prevalent model [29-32,34-38,40,42-46,49,51,53,54,56,57,60-64,67].

Figure 4. Characteristics of remote strategies in primary health care.



Impacts on Quality of Care

Studies reported the impacts of remote strategies on technical, interpersonal, or organizational dimensions of quality of care. Positive impacts were highlighted in 19 of the 44 (43%) studies [31,33,34,38,41,42,45,47-50, 52,56,57,60,62,64,65,70], negative impacts were mentioned in 6 (14%) studies [36,40,51,53,63,69], and positive and negative impacts were mentioned in 19 (43%) studies [29, 30, 32, 35, 37, 39, 43, 44, 46, 54, 55, 58, 59, 61, 66-68, 71, 72].

Technical [29-31,33-40,42-44,46,47,49-69,71,72] and organizational [29-43,45-50,52-61,64,65,67-72] dimensions were the most cited, followed by the interpersonal dimension [31,35-43,46-48,52,53,55,59,64,65,67-70,72]. More than one dimension of quality of care was directly or indirectly addressed in most studies.

Textbox 1 summarizes the positive and negative impacts on dimensions of quality of care evidenced in the studies.

Textbox 1. Impacts on dimensions of quality of care.**Technical dimension***Positive impacts*

- Security in care provision [31,33,35,37,38,42,43,46,49,52,55,57-59,61,64,66,67,71,72]
- Technical quality of information and communication technology [31,34,49,52,57,60,62,68]
- Technical accuracy [42,44,47,49,52,58,67]
- Resolvability [33,35,43,55,64,72]
- Support for clinical decision-making [50,65]
- Reliability [64,68]
- Utility [30]
- Attendance [56]
- Privacy [31]

Negative impacts

- Technical inaccuracy/inaccuracy [36,37,46,53,55,59,61,69,72]
- Low quality of consultations [33,36,51,63,69]
- Lack of assessment of vital signs and physical exams [29,59,63,69]
- Selective resolvability [39,40,53,54]
- Insecurity of data privacy [36,54,61,67]
- Discrepancy between professional conduct [66]

Interpersonal dimension*Positive impacts*

- Trust and bond with professionals improved adherence [43,47,48,65,67,68,70,72]
- Ease loneliness [39,46,52,59,64,70,72]
- Professional respect [31,42,43,64,68]
- Active listening [38,47,59]
- Positive interpersonal communication [38,41,64]
- Humanization of care [31]

Negative impacts

- Loss of nonverbal communication; lack of eye contact or touch [36,43,53,55,59,68,69,72]
- Interpersonal communication hampered by technology, speed of consultation, or memory difficulties of patients [39,40,53,59,69]
- Great emotional burden and stress [40,55]
- Fear of not being resolute compared with face-to-face modality; insecurity [35,37]

Organizational domain*Positive impacts*

- Continuous care [32-34,38,39,42,43,47,48,54-57,59,60,65,67,68,70]
- Economic, social, geographical, time, and cultural accessibility [29,31,34,38,42,52,54,55,57,67,70-72]
- Coordination of care [31,33,34,47,49,50,52,57-59,64]
- Access [42,43,45-47,54,58,65,72]
- Integrity of care [31,38,42,57,67,71]
- Optimization of consultation time [15,38,43,52,55,61,64]
- Economic efficiency [33,38,52,57,64]
- Organization of the work process [31,41,64]

- Increased demand for assistance [38,64,72]
- Planning of quality improvement [31,41]
- User-friendly technologies [33]
- Community engagement [37]

Negative impacts

- Reduced access; evidence of inequity [32,35-37,39,40,54,55,59,69]
- Reduced integrality of care [36,37,40,46,53,54,59,69]
- Digital exclusion [35-37,39,40,54,68,69]
- Lack of planning in defining the role of the team; disarticulation between actions and needs of the population [37,46,53,55,59,69,71]
- Reduced training of professionals using information and communication technology [37,53,54,59,61,69]
- Reduced continuity of care [30,36,40,54,58]
- Reduced coordination of care; fragile remote-presential articulation [37,46,54,55,68]
- Lack of professionals; high turnover [37,54,59]
- Reduced accessibility [30,32,53]
- Lack of support in internet technologies [36,69,71]
- Reduced active search in the community [53]

Benefits, Limitations, and Challenges of Digital Strategies in PHC

The following benefits of digital strategies in PHC were highlighted in the studies: (1) acceptability and patient satisfaction [29,31,34,36,38,43-45,47,52,54,56-58,62-64,67,68]; (2) great possibility of sustainability in the postpandemic period [31,32,38,40,43,54,55,58,62,64,68-72]; (3) increased frequency of people seeking care in PHC, especially in remote areas with difficult access and little face-to-face demand [29,34-36,41,46,49,50,53, 55,59,70-72]; (4) great safety against COVID-19 transmission [33,35,38,42,47,54, 55,57-59,61,66,71,72]; (5) time and cost savings due to geographic displacements [33,34,36,38, 43,49,55,63,68,70,72]; (6) organization of work process and scheduling of face-to-face and remote demands [31,39,41,47,48,53,54,63]; (7) faster service [33,52-54,61,66-68]; (8) reduced need for referrals to secondary care and hospitalizations [33,35,44,50,52,57,65]; (9) great comfort and practicality [34,36,42,43,68,72]; (10) optimization of training, meetings, and education of professionals [31,33,41,49,59]; (11) opportunity to be present in patients' lives, which benefits emotional health [30,32,43,44]; (12) fast home screening in cases of clinical changes [31,44,57]; (13) better communication with patients [46,64]; (14) great facility of use of technological tools and opportunity to overcome technological limitations [52,64]; (15) advantage of video calls over other tools [39,63]; (16) possibility of choosing the attendance modality [54]; (17) anonymity in situations that generate stigma, such as abortion care [42]; and (18) increased possibility of contacting inaccessible patients [61].

Conversely, the following limitations and challenges of digital strategies in PHC were identified: (1) difficulty in accessing internet, poor connectivity, digital divide (ie, more people with access to telephones and less to video calls) or digital desert (ie, people without access to technologies)

[32-36,38-40,42,45-47,54,55,59,64,67,68,72]; (2) increased need for training professionals and the population regarding digital health [36,37,39-41,44-47,50,51,55,59,61,66,69-71]; (3) great diagnostic imprecision and professional misconduct due to absence of physical examinations [39,43,52, 54-56,59-61,65,68,69,72]; (4) inconsistent platforms, with errors in data storage, limited resources, or both [31,33,38,41,43,45,58,60,64,68,71]; (5) difficult communication with the elderly, children, and people with disabilities or dementia [37-39,46,48,53,55,59,69]; (6) lack of planning regarding management of services [40,41,46,52,54,55,61,71]; (7) uncertainty about privacy and confidentiality of personal data [35,36,41,61,63,66,67]; (8) rapid implementation of remote services without prior guarantee of equitable access [30,42,55,63,71,72]; (9) poor support from information technology professionals [31,36,41,43,66,71]; (10) great need for good articulation between remote and face-to-face modalities to meet demands [39,40,60,63,70]; (11) mental stress in health workers [37,43,46,55,59]; (12) lack of health professionals, high turnover of professionals, or both [37,54,57,59,67]; (13) possible increase of chronic conditions (eg, certain groups of people who stopped seeking services) and side effects due to excessive self-medication [53,55,58,59]; (14) telephone calls are used but not resolute [34,35,53,64]; (15) low acceptability of professionals toward new remote workflows [46,51,55]; (16) difficult clinical monitoring of patients at home [51,57,64]; (17) difficulty regarding early identification of more complex health demands [31,59,69]; (18) delayed administrative tasks of health teams due to increased care demands [47,59]; (19) fast and urgent care [53,54]; (20) difficult articulation between professionals to meet more complex demands [44,54]; (21) difficulty regarding referral to other services [46]; (22) poor resolution in situations of risk at home (ie, domestic violence) [72]; (23) reduced supply of services [32]; and (24) difficulty in long-term follow-up of patients [49].

Discussion

Main Findings and Relation to Existing Literature

This scoping review demonstrated that the COVID-19 pandemic impacted health care in PHC worldwide (ie, fast implementation or increased use of remote care strategies or both) to mitigate the pandemic and ensure continuity of activities [73]. Various terms to refer to remote strategies were found in the literature [8,74]. Beyond concepts, technologies and tools are important components for health care systems, supporting the interaction among health care professionals or between health care professionals and patients [9]. The WHO [13,14] suggests telemedicine or telehealth to define distance care using ICT, whose purpose is to provide health care services in situations where distance or geographic barriers hinder the provision of care. Recently, “digital health” was introduced as an umbrella term, covering the use of electronic and mobile technologies (eg, advanced computer science, artificial intelligence, and big data) to support health and emerging health care areas [75].

The WHO and others [75,76] highlight the importance of digital technologies for achieving sustainable development goals and the advance of universal health coverage as opportunities to face challenges of health systems (ie, delayed provision of care, and reduced demand, adherence, and geographic accessibility) and increase coverage, accessibility, and quality of actions.

Telephone and video consultations are efficient tools for offering digital health [77,78]. Although telephone may increase follow-up contact and is more accessible than tools that require an internet connection, the assessment of severity and health status is compromised due to the absence of eye contact [79].

Telephone and audio consultations were recognized as telehealth modalities during the COVID-19 pandemic to support social distancing [80]. Although video consultations were rare in many locations before the pandemic [77], they are superior to phone calls, mainly due to eye contact and better communication for building bonds. Nevertheless, technical problems are more frequent when using digital strategies, and people need a stable connection to the internet, which may raise questions about the relationship between equity and the type of technology used [81-84]. For greater benefits, the literature indicates that the use of technology should be simple, consistent with local workflows, convenient for users, offer advantages over face-to-face consultations [76,85,86], and complement other existing technologies.

Results of this study corroborate with those of Breton et al [87], in which phone calls and video calls were identified as the most frequently used remote technologies, especially in the first months of the pandemic. We highlight that communication between health services users and professionals, mainly regarding platforms that ensure safety and reliability in the context of health care [88], is an important measure to be adopted due to the increased offer of newly developed applications.

The results of this scoping review also revealed the positive and negative impacts of remote strategies on quality of care in PHC worldwide, suggesting different types of organization (eg, fast

or urgent implementation) of digital strategies. Safe offer of care, technical quality and accuracy, and resolvability were the positive impacts most frequently reported in the technical dimension. By contrast, technical inaccuracy or imprecision, consultations with poor quality, lack of detailed physical examination, and selective solving of problems were also observed.

The interpersonal dimension was characterized by trust and bond with professionals that facilitated adherence to technologies, increased the possibility of talking to someone, alleviated loneliness caused by isolation, and improved respect between professionals and patients. From another perspective, we also found loss of nonverbal communication, lack of physical contact, difficult communication aggravated by technologies, and negative and stressful emotional load among professionals as negative impacts.

The impacts in the organizational dimension were the most frequently identified in the included studies, which strengthened continuity of care; economic, social, geographical, time, and cultural accessibility; coordination of care; access; integrity of care; and optimization of appointment time and efficiency. Negative impacts were also observed in this dimension, such as reduced access to services, inequity, and unequal use of services offered; digital exclusion of part of the population due to lack of technologies, connectivity, or knowledge regarding use; reduced integrity of care; lack of planning for defining the role of professionals; disarticulation of actions with real needs of the population; impaired continuity of care; reduced coordination of care; fragile articulation between remote and face-to-face modalities; and unpreparedness of professionals to meet demands mediated by ICT.

One study [89] that verified how the pandemic impacted primary care services suggested digital health as an inflection point for PHC and the only alternative for restructuring the workflow of health care providers during the pandemic. The latter may have also contributed to the impaired quality of health care, especially for the elderly and people with preexisting health conditions (ie, psychological problems, addictions, or victims of domestic violence).

Issues limiting technological barriers and ethics in the use of information might be linked to work organization, health financing, and lack of familiarity of professionals and patients [6]. When properly available, patients considered digital health to be satisfactory and safe, and felt comfortable when trusting relationships with professionals and person-centered practices were present.

In PHC, preexisting virtual solutions to COVID-19 served as opportunities to support public health responses in combating the pandemic and minimizing the risk of exposure [90-93]. The adaptation of health systems based on PHC and training of professionals regarding the use of digital tools to fulfill clinical responsibilities, which previously required face-to-face contact, were also useful [90]. Studies also highlighted the relevance of digital strategies in preventive and health promotion actions, such as remote monitoring of clinical signs; management of chronic diseases and medication; and guidance on healthy lifestyle, exercises, and eating habits [94,95].

Studies conducted before the COVID-19 pandemic demonstrated the importance of digital health in expanding access in PHC [82,96,97], even though face-to-face care was preferred [98]. Positive experiences were associated with planning according to the health needs of the population [99-101], whereas health professionals complained about insufficient remuneration, unavailability of technologies, and lack of standardization [102,103]. Based on these prepandemic experiences, digital strategies in PHC were an option to mitigate barriers and increase access for hard-to-reach populations. During periods of greater restriction and social isolation due to the COVID-19 pandemic, the reality of virtual assistance was extrapolated beyond populations with difficulties in accessing services. This fact allowed us to observe different results regarding the strengthening of digital health or predominance of persistent problems that depended on decision-making factors of governance to provide broad coverage of technologies (complementary or alternative) to populations. In fact, in most situations, digital health was adopted without the support of a national strategy.

The results of this study emphasize the benefits, limitations, and challenges of remote strategies in PHC, offering lessons during a global public health crisis. In this sense, quality of care in PHC can still be improved with consolidation and advances in digital health.

Implications for Practice and Research

According to the Pan American Health Organization [104], ICTs are essential to increase access of citizens to high-quality PHC, regardless of their distance from large urban centers. Technologies are becoming the primary method in which people, governments, and health institutions work, communicate, and generate and exchange knowledge. In this context, we must reflect on how remote technologies and strategies can support and strengthen essential characteristics of PHC, since this is the first point of contact for people, and offers comprehensive, accessible, and community-based health care. PHC also offers health promotion and prevention, treatment of acute and infectious diseases, control of chronic diseases, palliative care, and rehabilitation to individuals, families, and communities [105].

This study demonstrates that the fast transition and expansion of digital health impacted access and quality of care in PHC worldwide, even considering that health needs, policies, management, and financing differ between countries. PHC must take advantage of the lessons learned from the COVID-19 pandemic, strengthen its response capacity, balance the offer of new modalities of care with expanded use of technologies, and be more equitable and accessible. In contrast, equity of health care supply is beyond the power of action of health professionals or management of local services, since it is a larger and structural problem that depends on the integrated actions and engagement of public and social policies.

PHC services must be aligned with the needs and satisfaction of the population, while efforts must be made to perform self-assessments and improve quality of in-person and remote processes. Planning and intersectoral articulation at the management level, along with investment in financial and human

resources are essential to improve the cost-effectiveness of remote care. Furthermore, technical and operational infrastructure is imperative for using technologies, strengthening security and protection of the patients and professional data.

Services and actions exceeding needs increase costs and do not improve results regarding patient-centered care and needs [106,107]. Moreover, health outcomes are worse, and costs are high when care is not based on the needs of the population. For digital health strategies in PHC, Lillrank et al [108] recommend planning actions by homogeneous groups with similar health needs, and organizing the supply of care considering demand, severity, and duration of needs, according to demand and supply-based operating modes. This organization could also facilitate continuity of care and optimize the work process using remote strategies.

The identification of gaps in the literature is expected in scoping reviews. As the COVID-19 pandemic changed the provision of services at all levels of care worldwide (eg, expansion of remote care strategies), directions for future research are challenging because the long-term impacts are unknown. Based on the observations from this scoping review, we recommend the following primary studies focused on remote strategies in PHC, especially in countries that have not yet investigated the topics discussed here: (1) assess implementation and differences between health systems (either public and private or with different forms of management and financing) based on the principles of universality and universal coverage; (2) assess the effectiveness and safety of remote strategies between users, professionals, and health managers; (3) monitor the impacts of remote strategies on quality of care and investigate how to enhance quality; and (4) perform intervention studies to investigate innovative strategies or approaches to improve clinical practice. Moreover, systematic reviews with meta-analysis could be performed to (1) assess the impact of remote strategies on clinical outcomes in vulnerable populations and (2) follow-up of patients with COVID-19 complications using ICT.

Consultation With Stakeholders

In the consultation stage, stakeholders were asked about ideas for future research, applicability of the results, and dissemination strategies. From the perspective of the participants, this scoping review can stimulate development agencies to finance ICT in PHC; reflect on cost-effectiveness of digital health to achieve greater adherence to therapeutic plans, reduce disease transmission, and prevent injuries; demonstrate the benefits of using digital health for monitoring indicators, goals, and indices in PHC, and for health surveillance; and support health professionals with lessons learned for improving care in remote mode.

Regarding the possibilities of disseminating the results, the following suggestions were discussed: scientific dissemination (indexed journals, conferences, and workshops); disclosure by health secretariats; creation of networks with interested social agents; linking of agents to research groups to approximate academia from health services and the general population; meetings and debates with local and national health managers; and adaptation of dissemination of results according to the local

culture, choosing the most accessible means of communication (ie, social networks).

When asked about ideas for future research, the following were suggested: action research with health managers and professionals focusing on solutions for digital inclusion of vulnerable populations; sectorial studies inserted in PHC (eg, sectional and intervention research designs regarding digital pharmaceutical and oral health care industries); studies investigating the acceptability of remote strategies by specific groups and its associated factors (eg, age, gender, socioeconomic status, preexisting health conditions, and beliefs); and long-term follow-up of patients using remote monitoring in PHC.

Strengths and Limitations

This scoping review is the first to broadly map evidence regarding the use of remote strategies in PHC and its impacts on quality of care in the context of COVID-19. The study met the criteria for scoping reviews [24,109], and followed methodological references, checklists, and published protocols [23].

We did not conduct a meta-analysis [23] or assess the quality of studies. However, these steps are not essential due to the

exploratory and descriptive nature of a scoping review. The search was performed to retrieve the highest number of publications regarding the topic, rather than focusing on studies with the highest standards of scientific rigor. Even though databases for peer-reviewed publications and gray literature were included with no filter limits and a high-sensitivity search strategy was performed, we do not know to what extent relevant studies and important databases were included.

Conclusion

This review provides information on the use of digital strategies in PHC and its impacts on quality of care during the COVID-19 pandemic. Confronting a public health situation of such proportion sheds light on realities that were not as evident previously. Given the importance of digital health in the current global health situation and the possibility of integrating and advancing this strategy after the pandemic, primary care must strengthen its response capacity, expand ICT use, and manage challenges using scientific evidence.

The number of digital health initiatives launched worldwide without a scientific basis during the pandemic had its foundation in the health crisis. Digital health needs to be improved and expanded to strengthen primary care and health systems.

Acknowledgments

The authors acknowledge the support provided by the Graduate Program in Health Sciences of the Federal University of Rio Grande do Norte and the Coordenação de Aperfeiçoamento de Pessoal de Nível Superior-Brazil (CAPES) (finance code 001). The authors thank Probatius Academic Services for providing scientific language translation, revision, and editing.

Authors' Contributions

SAdCU, CRDVS, RHL, and MF-T planned the study. CRDVS and RHL performed article selection and data extraction, and SAdCU was the third reviewer. CRDVS, RHL, OdGBJ, and CSM performed the analysis and synthesis of results. CRDVS conducted the consultation with stakeholders. All authors contributed to writing the manuscript. SAdCU, MF-T, RAA, LVL, and SD critically reviewed the manuscript. All authors approved the final version of the manuscript.

Conflicts of Interest

None declared.

Multimedia Appendix 1

Search strategy.

[[DOC File , 47 KB - humanfactors_v9i2e35380_app1.doc](#)]

Multimedia Appendix 2

Data extraction spreadsheet (original language).

[[XLSX File \(Microsoft Excel File\), 127 KB - humanfactors_v9i2e35380_app2.xlsx](#)]

Multimedia Appendix 3

Stakeholders consultation guide.

[[DOC File , 26 KB - humanfactors_v9i2e35380_app3.doc](#)]

References

1. Committee on Quality of Health Care in America, Institute of Medicine. Crossing the quality chasm: a new health system for the 21st century. Washington, DC: National Academy Press; 2001.
2. Donabedian A. Epigenomics. In: An introduction to quality assurance in health care. New York: Oxford University Press; 2002.

3. Champagne F, Hartz Z, Brousselle A. A apreciação normativa. In: Brousselle A, Champagne F, Contradriopoulos AP, Hartz Z, editors. Avaliação: conceitos e métodos. Rio de Janeiro, Brazil: Fiocruz; 2011:77-94.
4. Donabedian A. Explorations in quality assessment and monitoring: the definition of quality and approaches to its assessment. Ann Arbor, MI: Health Administration Press; 1980.
5. Rawaf S, Allen LN, Stigler FL, Kringos D, Quezada Yamamoto H, van Weel C, Global Forum on Universal Health Coverage Primary Health Care. Lessons on the COVID-19 pandemic, for and by primary care professionals worldwide. Eur J Gen Pract 2020 Dec;26(1):129-133 [FREE Full text] [doi: [10.1080/13814788.2020.1820479](https://doi.org/10.1080/13814788.2020.1820479)] [Medline: [32985278](https://pubmed.ncbi.nlm.nih.gov/32985278/)]
6. Garattini L, Badinella Martini M, Mannucci PM. Improving primary care in Europe beyond COVID-19: from telemedicine to organizational reforms. Intern Emerg Med 2021 Mar;16(2):255-258 [FREE Full text] [doi: [10.1007/s11739-020-02559-x](https://doi.org/10.1007/s11739-020-02559-x)] [Medline: [33196973](https://pubmed.ncbi.nlm.nih.gov/33196973/)]
7. Liao C, Chang W, Yu W, Toh HS. Management of acute cardiovascular events in patients with COVID-19. Rev Cardiovasc Med 2020 Dec 30;21(4):577-581 [FREE Full text] [doi: [10.31083/j.rcm.2020.04.140](https://doi.org/10.31083/j.rcm.2020.04.140)] [Medline: [33388002](https://pubmed.ncbi.nlm.nih.gov/33388002/)]
8. Waqas A, Teoh SH, Lapão LV, Messina LA, Correia JC. Harnessing telemedicine for the provision of health care: bibliometric and scientometric analysis. J Med Internet Res 2020 Oct 02;22(10):e18835 [FREE Full text] [doi: [10.2196/18835](https://doi.org/10.2196/18835)] [Medline: [33006571](https://pubmed.ncbi.nlm.nih.gov/33006571/)]
9. Tuckson RV, Edmunds M, Hodgkins ML. Telehealth. N Engl J Med 2017 Oct 19;377(16):1585-1592. [doi: [10.1056/nejmsr1503323](https://doi.org/10.1056/nejmsr1503323)]
10. Stefos T, Carey K, Shen M, Poe S, Oh D, Moran E. The effect of telehealth services on provider productivity. Med Care 2021 May 01;59(5):456-460. [doi: [10.1097/MLR.0000000000001529](https://doi.org/10.1097/MLR.0000000000001529)] [Medline: [33821831](https://pubmed.ncbi.nlm.nih.gov/33821831/)]
11. Snoswell CL, Taylor ML, Comans TA, Smith AC, Gray LC, Caffery LJ. Determining if telehealth can reduce health system costs: scoping review. J Med Internet Res 2020 Oct 19;22(10):e17298 [FREE Full text] [doi: [10.2196/17298](https://doi.org/10.2196/17298)] [Medline: [33074157](https://pubmed.ncbi.nlm.nih.gov/33074157/)]
12. Wosik J, Fudim M, Cameron B, Gellad ZF, Cho A, Phinney D, et al. Telehealth transformation: COVID-19 and the rise of virtual care. J Am Med Inform Assoc 2020 Jun 01;27(6):957-962 [FREE Full text] [doi: [10.1093/jamia/ocaa067](https://doi.org/10.1093/jamia/ocaa067)] [Medline: [32311034](https://pubmed.ncbi.nlm.nih.gov/32311034/)]
13. WHO Global Observatory for eHealth. Telemedicine: opportunities and developments in Member States: report on the second global survey on eHealth. Geneva: World Health Organization; 2010.
14. A health telematics policy in support of WHO's Health-for-all strategy for global health development: report of the WHO Group Consultation on Health Telematics, 11-16 December, Geneva, 1997. World Health Organization, Institutional Repository for Information Sharing. 1997. URL: <https://apps.who.int/iris/handle/10665/63857> [accessed 2021-11-20]
15. Estrategia mundial sobre salud digital 2020–2025. World Health Organization, Institutional Repository for Information Sharing. 2020. URL: https://www.paho.org/hq/index.php?option=com_docman&view=download&category_slug=56-directing-council-spanish- [accessed 2021-11-20]
16. Mann DM, Chen J, Chunara R, Testa PA, Nov O. COVID-19 transforms health care through telemedicine: Evidence from the field. J Am Med Inform Assoc 2020 Jul 01;27(7):1132-1135 [FREE Full text] [doi: [10.1093/jamia/ocaa072](https://doi.org/10.1093/jamia/ocaa072)] [Medline: [32324855](https://pubmed.ncbi.nlm.nih.gov/32324855/)]
17. Bhaskar S, Bradley S, Sakhamuri S, Moguilner S, Chattu VK, Pandya S, et al. Designing futuristic telemedicine using artificial intelligence and robotics in the COVID-19 era. Front Public Health 2020 Nov 2;8:556789. [doi: [10.3389/fpubh.2020.556789](https://doi.org/10.3389/fpubh.2020.556789)] [Medline: [33224912](https://pubmed.ncbi.nlm.nih.gov/33224912/)]
18. Webster P. Virtual health care in the era of COVID-19. Lancet 2020 Apr 11;395(10231):1180-1181 [FREE Full text] [doi: [10.1016/S0140-6736\(20\)30818-7](https://doi.org/10.1016/S0140-6736(20)30818-7)] [Medline: [32278374](https://pubmed.ncbi.nlm.nih.gov/32278374/)]
19. Peters M, Godfrey C, McInerney P, Munn Z, Tricco A, Khalil H. Chapter 11: scoping reviews. In: Aromataris E, Munn Z, editors. JBI manual for evidence synthesis. Adelaide, Australia: JBI Collaboration; 2020.
20. Tricco AC, Lillie E, Zarin W, O'Brien KK, Colquhoun H, Levac D, et al. PRISMA Extension for Scoping Reviews (PRISMA-ScR): Checklist and Explanation. Ann Intern Med 2018 Oct 02;169(7):467-473 [FREE Full text] [doi: [10.7326/M18-0850](https://doi.org/10.7326/M18-0850)] [Medline: [30178033](https://pubmed.ncbi.nlm.nih.gov/30178033/)]
21. Arksey H, O'Malley L. Scoping studies: towards a methodological framework. Int J Soc Res Method 2005 Feb;8(1):19-32. [doi: [10.1080/1364557032000119616](https://doi.org/10.1080/1364557032000119616)]
22. Levac D, Colquhoun H, O'Brien KK. Scoping studies: advancing the methodology. Implement Sci 2010 Sep 20;5:69 [FREE Full text] [doi: [10.1186/1748-5908-5-69](https://doi.org/10.1186/1748-5908-5-69)] [Medline: [20854677](https://pubmed.ncbi.nlm.nih.gov/20854677/)]
23. Silva CRDV, Lopes RH, de Goes Bay Júnior O, Fuentealba-Torres M, Arcêncio RA, da Costa Uchôa SA. Telemedicine in primary healthcare for the quality of care in times of COVID-19: a protocol. BMJ Open 2021 Jul 12;11(7):e046227 [FREE Full text] [doi: [10.1136/bmjopen-2020-046227](https://doi.org/10.1136/bmjopen-2020-046227)] [Medline: [34253666](https://pubmed.ncbi.nlm.nih.gov/34253666/)]
24. Peters MDJ, Marnie C, Tricco AC, Pollock D, Munn Z, Alexander L, et al. Updated methodological guidance for the conduct of scoping reviews. JBI Evid Synth 2020 Oct;18(10):2119-2126. [doi: [10.11124/JBIES-20-00167](https://doi.org/10.11124/JBIES-20-00167)] [Medline: [33038124](https://pubmed.ncbi.nlm.nih.gov/33038124/)]
25. Oliveira Araújo WC. Recuperação da informação em saúde. ConCI 2020 Jul 10;3(2):100-134. [doi: [10.33467/conci.v3i2.13447](https://doi.org/10.33467/conci.v3i2.13447)]

26. Shamseer L, Moher D, Clarke M, Ghersi D, Liberati A, Petticrew M, PRISMA-P Group. Preferred reporting items for systematic review and meta-analysis protocols (PRISMA-P) 2015: elaboration and explanation. *BMJ* 2015 Jan 02;350:g7647 [FREE Full text] [doi: [10.1136/bmj.g7647](https://doi.org/10.1136/bmj.g7647)] [Medline: [25555855](https://pubmed.ncbi.nlm.nih.gov/25555855/)]
27. Ouzzani M, Hammady H, Fedorowicz Z, Elmagarmid A. Rayyan-a web and mobile app for systematic reviews. *Syst Rev* 2016 Dec 05;5(1):210 [FREE Full text] [doi: [10.1186/s13643-016-0384-4](https://doi.org/10.1186/s13643-016-0384-4)] [Medline: [27919275](https://pubmed.ncbi.nlm.nih.gov/27919275/)]
28. Braun V, Clarke V. Using thematic analysis in psychology. *Qual Res Psychol* 2006 Jan;3(2):77-101. [doi: [10.1191/1478088706qp063oa](https://doi.org/10.1191/1478088706qp063oa)]
29. Alexander GC, Tajanlangit M, Heyward J, Mansour O, Qato DM, Stafford RS. Use and content of primary care office-based vs telemedicine care visits during the COVID-19 pandemic in the US. *JAMA Netw Open* 2020 Oct 01;3(10):e2021476 [FREE Full text] [doi: [10.1001/jamanetworkopen.2020.21476](https://doi.org/10.1001/jamanetworkopen.2020.21476)] [Medline: [33006622](https://pubmed.ncbi.nlm.nih.gov/33006622/)]
30. Schweiberger K, Hoberman A, Iagnemma J, Schoemer P, Squire J, Taormina J, et al. Practice-level variation in telemedicine use in a pediatric primary care network during the COVID-19 pandemic: retrospective analysis and survey study. *J Med Internet Res* 2020 Dec 18;22(12):e24345 [FREE Full text] [doi: [10.2196/24345](https://doi.org/10.2196/24345)] [Medline: [33290244](https://pubmed.ncbi.nlm.nih.gov/33290244/)]
31. Olayiwola JN, Magaña C, Harmon A, Nair S, Esposito E, Harsh C, et al. Telehealth as a bright spot of the COVID-19 pandemic: recommendations from the virtual frontlines ("Frontweb"). *JMIR Public Health Surveill* 2020 Jun 25;6(2):e19045 [FREE Full text] [doi: [10.2196/19045](https://doi.org/10.2196/19045)] [Medline: [32479413](https://pubmed.ncbi.nlm.nih.gov/32479413/)]
32. Atherly A, Van Den Broek-Altensburg E, Hart V, Gleason K, Carney J. Consumer reported care deferrals due to the COVID-19 pandemic, and the role and potential of telemedicine: cross-sectional analysis. *JMIR Public Health Surveill* 2020 Sep 14;6(3):e21607 [FREE Full text] [doi: [10.2196/21607](https://doi.org/10.2196/21607)] [Medline: [32833661](https://pubmed.ncbi.nlm.nih.gov/32833661/)]
33. Judson TJ, Odisho AY, Neinstein AB, Chao J, Williams A, Miller C, et al. Rapid design and implementation of an integrated patient self-triage and self-scheduling tool for COVID-19. *J Am Med Inform Assoc* 2020 Jun 01;27(6):860-866 [FREE Full text] [doi: [10.1093/jamia/ocaa051](https://doi.org/10.1093/jamia/ocaa051)] [Medline: [32267928](https://pubmed.ncbi.nlm.nih.gov/32267928/)]
34. Mills KT, Peacock E, Chen J, Zimmerman A, He H, Cyprian A, et al. Experiences and beliefs of low-income patients with hypertension in Louisiana and Mississippi during the COVID-19 pandemic. *J Am Heart Assoc* 2021 Feb 02;10(3):e018510 [FREE Full text] [doi: [10.1161/JAHA.120.018510](https://doi.org/10.1161/JAHA.120.018510)] [Medline: [33267723](https://pubmed.ncbi.nlm.nih.gov/33267723/)]
35. Tarn DM, Hintz C, Mendez-Hernandez E, Sawlani SP, Bholat MA. Using virtual visits to care for primary care patients with COVID-19 symptoms. *J Am Board Fam Med* 2021 Feb 23;34(Supplement):S147-S151. [doi: [10.3122/jabfm.2021.s1.200241](https://doi.org/10.3122/jabfm.2021.s1.200241)]
36. Adepoju O, Liaw W, Chae M, Ojinnaka C, Britton E, Reves S, et al. COVID-19 and telehealth operations in Texas primary care clinics: disparities in medically underserved area clinics. *J Health Care Poor Underserved* 2021;32(2):948-957. [doi: [10.1353/hpu.2021.0073](https://doi.org/10.1353/hpu.2021.0073)] [Medline: [34120986](https://pubmed.ncbi.nlm.nih.gov/34120986/)]
37. Ritchie CS, Gallopyn N, Sheehan OC, Sharieff SA, Franzosa E, Gorbenko K, et al. COVID challenges and adaptations among home-based primary care practices: lessons for an ongoing pandemic from a national survey. *J Am Med Dir Assoc* 2021 Jul;22(7):1338-1344 [FREE Full text] [doi: [10.1016/j.jamda.2021.05.016](https://doi.org/10.1016/j.jamda.2021.05.016)] [Medline: [34111388](https://pubmed.ncbi.nlm.nih.gov/34111388/)]
38. Drerup B, Espenschied J, Wiedemer J, Hamilton L. Reduced no-show rates and sustained patient satisfaction of telehealth during the COVID-19 pandemic. *Telemed J E Health* 2021 Dec;27(12):1409-1415. [doi: [10.1089/tmj.2021.0002](https://doi.org/10.1089/tmj.2021.0002)] [Medline: [33661708](https://pubmed.ncbi.nlm.nih.gov/33661708/)]
39. Kalicki AV, Moody KA, Franzosa E, Gliatto PM, Ornstein KA. Barriers to telehealth access among homebound older adults. *J Am Geriatr Soc* 2021 Sep;69(9):2404-2411 [FREE Full text] [doi: [10.1111/jgs.17163](https://doi.org/10.1111/jgs.17163)] [Medline: [33848360](https://pubmed.ncbi.nlm.nih.gov/33848360/)]
40. Chang JE, Lai AY, Gupta A, Nguyen AM, Berry CA, Shelley DR. Rapid transition to telehealth and the digital divide: implications for primary care access and equity in a post-COVID era. *Milbank Q* 2021 Jun;99(2):340-368 [FREE Full text] [doi: [10.1111/1468-0009.12509](https://doi.org/10.1111/1468-0009.12509)] [Medline: [34075622](https://pubmed.ncbi.nlm.nih.gov/34075622/)]
41. Thies KM, Gonzalez M, Porto A, Ashley KL, Korman S, Lamb M. Project ECHO COVID-19: vulnerable populations and telehealth early in the pandemic. *J Prim Care Community Health* 2021;12:21501327211019286 [FREE Full text] [doi: [10.1177/21501327211019286](https://doi.org/10.1177/21501327211019286)] [Medline: [34036832](https://pubmed.ncbi.nlm.nih.gov/34036832/)]
42. Godfrey EM, Thayer EK, Fiaastro AE, Aiken ARA, Gomperts R. Family medicine provision of online medication abortion in three US states during COVID-19. *Contraception* 2021 Jul;104(1):54-60 [FREE Full text] [doi: [10.1016/j.contraception.2021.04.026](https://doi.org/10.1016/j.contraception.2021.04.026)] [Medline: [33939985](https://pubmed.ncbi.nlm.nih.gov/33939985/)]
43. Juarez-Reyes M, Mui HZ, Kling SM, Brown-Johnson C. Accessing behavioral health care during COVID: rapid transition from in-person to teleconferencing medical group visits. *Ther Adv Chronic Dis* 2021 Feb 12;12:2040622321990269 [FREE Full text] [doi: [10.1177/2040622321990269](https://doi.org/10.1177/2040622321990269)] [Medline: [33633823](https://pubmed.ncbi.nlm.nih.gov/33633823/)]
44. Bui V, Ackerman S. Hamilton Primary Care Virtual Care Survey Analysis. Ontario Ministry of Health and Long-term Care. URL: http://health.gov.on.ca/en/pro/programs/publichealth/coronavirus/docs/2019_primary_care_guidance.pdf [accessed 2021-11-20]
45. Mohammed HT, Hyseni L, Bui V, Gerritsen B, Fuller K, Sung J, et al. Exploring the use and challenges of implementing virtual visits during COVID-19 in primary care and lessons for sustained use. *PLoS One* 2021;16(6):e0253665 [FREE Full text] [doi: [10.1371/journal.pone.0253665](https://doi.org/10.1371/journal.pone.0253665)] [Medline: [34166441](https://pubmed.ncbi.nlm.nih.gov/34166441/)]

46. Donnelly C, Ashcroft R, Bobbette N, Mills C, Mofina A, Tran T, et al. Interprofessional primary care during COVID-19: a survey of the provider perspective. *BMC Fam Pract* 2021 Feb 03;22(1):31 [FREE Full text] [doi: [10.1186/s12875-020-01366-9](https://doi.org/10.1186/s12875-020-01366-9)] [Medline: [33535973](https://pubmed.ncbi.nlm.nih.gov/33535973/)]
47. Araujo Gomes de Castro F, Oliveira dos Santos ?, Valadares Labanca Reis G, Brandão Viveiros L, Hespanhol Torres M, De Oliveira Junior PP. Telemedicina rural e COVID-19. *Rev Bras Med Fam Comunidade* 2020 Jun 24;15(42):2484. [doi: [10.5712/rbmfc15\(42\)2484](https://doi.org/10.5712/rbmfc15(42)2484)]
48. Dimer NA, Canto-Soares ND, Santos-Teixeira LD, Goulart BNGD. The COVID-19 pandemic and the implementation of telehealth in speech-language and hearing therapy for patients at home: an experience report. *Codas* 2020;32(3):e20200144 [FREE Full text] [doi: [10.1590/2317-1782/20192020144](https://doi.org/10.1590/2317-1782/20192020144)] [Medline: [32578694](https://pubmed.ncbi.nlm.nih.gov/32578694/)]
49. Queiroz MS, de Carvalho JX, Bortoto SF, de Matos MR, das Graças Dias Cavalcante C, Andrade EAS, et al. Diabetic retinopathy screening in urban primary care setting with a handheld smartphone-based retinal camera. *Acta Diabetol* 2020 Dec;57(12):1493-1499 [FREE Full text] [doi: [10.1007/s00592-020-01585-7](https://doi.org/10.1007/s00592-020-01585-7)] [Medline: [32748176](https://pubmed.ncbi.nlm.nih.gov/32748176/)]
50. Silva RSD, Schmitz CAA, Harzheim E, Molina-Bastos CG, Oliveira EBD, Roman R, et al. The role of telehealth in the Covid-19 pandemic: a Brazilian experience. *Cien Saude Colet* 2021 Jun;26(6):2149-2157 [FREE Full text] [doi: [10.1590/1413-81232021266.39662020](https://doi.org/10.1590/1413-81232021266.39662020)] [Medline: [34231727](https://pubmed.ncbi.nlm.nih.gov/34231727/)]
51. Sahni M, Choudhry J, Mittal A, Bhogal G. Remote musculoskeletal consultations: a survey of general practitioner registrars' level of confidence, acceptability, and management. *Cureus* 2021 May 18;13(5):e15084 [FREE Full text] [doi: [10.7759/cureus.15084](https://doi.org/10.7759/cureus.15084)] [Medline: [34150413](https://pubmed.ncbi.nlm.nih.gov/34150413/)]
52. Leung K, Qureshi S. Managing high frequency users of an electronic consultation system in primary care: a quality improvement project. *BMJ Open Qual* 2021 Jun;10(2):e001310 [FREE Full text] [doi: [10.1136/bmjopen-2020-001310](https://doi.org/10.1136/bmjopen-2020-001310)] [Medline: [34112657](https://pubmed.ncbi.nlm.nih.gov/34112657/)]
53. Tuijt R, Rait G, Frost R, Wilcock J, Manthorpe J, Walters K. Remote primary care consultations for people living with dementia during the COVID-19 pandemic: experiences of people living with dementia and their carers. *Br J Gen Pract* 2021 Aug;71(709):e574-e582 [FREE Full text] [doi: [10.3399/BJGP.2020.1094](https://doi.org/10.3399/BJGP.2020.1094)] [Medline: [33630749](https://pubmed.ncbi.nlm.nih.gov/33630749/)]
54. Salisbury C, Quigley A, Hex N, Aznar C. Private video consultation services and the future of primary care. *J Med Internet Res* 2020 Oct 01;22(10):e19415 [FREE Full text] [doi: [10.2196/19415](https://doi.org/10.2196/19415)] [Medline: [32812887](https://pubmed.ncbi.nlm.nih.gov/32812887/)]
55. Murphy M, Scott LJ, Salisbury C, Turner A, Scott A, Denholm R, et al. Implementation of remote consulting in UK primary care following the COVID-19 pandemic: a mixed-methods longitudinal study. *Br J Gen Pract* 2021;71(704):e166-e177 [FREE Full text] [doi: [10.3399/BJGP.2020.0948](https://doi.org/10.3399/BJGP.2020.0948)] [Medline: [33558332](https://pubmed.ncbi.nlm.nih.gov/33558332/)]
56. Saiz Llamosas J, Pérez García R. Eficacia del tratamiento fisioterápico en atención primaria, mediante consulta no presencial, a un paciente dado de alta de neumonía por Coronavirus. *Fisioterapia* 2021 Jan;43(1):58-62. [doi: [10.1016/j.ft.2020.09.003](https://doi.org/10.1016/j.ft.2020.09.003)]
57. Coronado-Vázquez V, Ramírez-Durán MDV, Gómez-Salgado J, Dorado-Rabaneda MS, Benito-Alonso E, Holgado-Juan M, et al. Evolution of a cohort of COVID-19 infection suspects followed-up from primary health care. *J Pers Med* 2021 May 24;11(6):459 [FREE Full text] [doi: [10.3390/jpm11060459](https://doi.org/10.3390/jpm11060459)] [Medline: [34073666](https://pubmed.ncbi.nlm.nih.gov/34073666/)]
58. Morreel S, Philips H, Verhoeven V. Organisation and characteristics of out-of-hours primary care during a COVID-19 outbreak: a real-time observational study. *PLoS One* 2020;15(8):e0237629 [FREE Full text] [doi: [10.1371/journal.pone.0237629](https://doi.org/10.1371/journal.pone.0237629)] [Medline: [32790804](https://pubmed.ncbi.nlm.nih.gov/32790804/)]
59. Verhoeven V, Tsakitzidis G, Philips H, Van Royen P. Impact of the COVID-19 pandemic on the core functions of primary care: will the cure be worse than the disease? A qualitative interview study in Flemish GPs. *BMJ Open* 2020 Jun 17;10(6):e039674 [FREE Full text] [doi: [10.1136/bmjopen-2020-039674](https://doi.org/10.1136/bmjopen-2020-039674)] [Medline: [32554730](https://pubmed.ncbi.nlm.nih.gov/32554730/)]
60. Johnsen TM, Norberg BL, Kristiansen E, Zanaboni P, Austad B, Krogh FH, et al. Suitability of video consultations during the COVID-19 pandemic lockdown: cross-sectional survey among Norwegian general practitioners. *J Med Internet Res* 2021 Feb 08;23(2):e26433 [FREE Full text] [doi: [10.2196/26433](https://doi.org/10.2196/26433)] [Medline: [33465037](https://pubmed.ncbi.nlm.nih.gov/33465037/)]
61. Lapão LV, Peyroteo M, Maia M, Seixas J, Gregório J, Mira da Silva M, et al. Implementation of digital monitoring services during the COVID-19 pandemic for patients with chronic diseases: design science approach. *J Med Internet Res* 2021 Aug 26;23(8):e24181 [FREE Full text] [doi: [10.2196/24181](https://doi.org/10.2196/24181)] [Medline: [34313591](https://pubmed.ncbi.nlm.nih.gov/34313591/)]
62. Florea M, Lazea C, Gaga R, Sur G, Lotrean L, Puia A, et al. Lights and shadows of the perception of the use of telemedicine by Romanian family doctors during the COVID-19 pandemic. *Int J Gen Med* 2021;14:1575-1587. [doi: [10.2147/IJGM.S309519](https://doi.org/10.2147/IJGM.S309519)] [Medline: [33953605](https://pubmed.ncbi.nlm.nih.gov/33953605/)]
63. Mueller M, Knop M, Niehaves B, Adarkwah CC. Investigating the acceptance of video consultation by patients in rural primary care: empirical comparison of preusers and actual users. *JMIR Med Inform* 2020 Oct 22;8(10):e20813 [FREE Full text] [doi: [10.2196/20813](https://doi.org/10.2196/20813)] [Medline: [32969339](https://pubmed.ncbi.nlm.nih.gov/32969339/)]
64. Kludacz-Alessandri M, Hawrysz L, Korneta P, Gierszewska G, Pomaranik W, Walczak R. The impact of medical teleconsultations on general practitioner-patient communication during COVID-19: a case study from Poland. *PLoS One* 2021;16(7):e0254960 [FREE Full text] [doi: [10.1371/journal.pone.0254960](https://doi.org/10.1371/journal.pone.0254960)] [Medline: [34270587](https://pubmed.ncbi.nlm.nih.gov/34270587/)]
65. Zeltzer D, Einav L, Rashba J, Balicer R. The impact of increased access to telemedicine. Stanford. URL: <http://web.stanford.edu/~leinav/wp/telemed.pdf> [accessed 2021-11-20]

66. Hasani SA, Ghafri TA, Al Lawati H, Mohammed J, Al Mukhainai A, Al Ajmi F, et al. The use of telephone consultation in primary health care during COVID-19 pandemic, Oman: perceptions from physicians. *J Prim Care Community Health* 2020;11:2150132720976480 [FREE Full text] [doi: [10.1177/2150132720976480](https://doi.org/10.1177/2150132720976480)] [Medline: [33307943](https://pubmed.ncbi.nlm.nih.gov/33307943/)]
67. Alharbi K, Aldosari M, Alhassan A, Alshallal K, Altamimi A, Altulaihi B. Patient satisfaction with virtual clinic during coronavirus disease (COVID-19) pandemic in primary healthcare, Riyadh, Saudi Arabia. *J Fam Community Med* 2021;28(1):48. [doi: [10.4103/jfcm.jfcm_353_20](https://doi.org/10.4103/jfcm.jfcm_353_20)]
68. Jannati N, Nakhaee N, Yazdi-Feyzabadi V, Tjondronegoro D. A cross-sectional online survey on patients' satisfaction using store-and-forward voice and text messaging teleconsultation service during the COVID-19 pandemic. *Int J Med Inform* 2021 Jul;151:104474 [FREE Full text] [doi: [10.1016/j.ijmedinf.2021.104474](https://doi.org/10.1016/j.ijmedinf.2021.104474)] [Medline: [33965682](https://pubmed.ncbi.nlm.nih.gov/33965682/)]
69. Isautier JM, Copp T, Ayre J, Cvejic E, Meyerowitz-Katz G, Batcup C, et al. People's experiences and satisfaction with telehealth during the COVID-19 pandemic in Australia: cross-sectional survey study. *J Med Internet Res* 2020 Dec 10;22(12):e24531 [FREE Full text] [doi: [10.2196/24531](https://doi.org/10.2196/24531)] [Medline: [33156806](https://pubmed.ncbi.nlm.nih.gov/33156806/)]
70. Javanparast S, Roeger L, Kwok Y, Reed RL. The experience of Australian general practice patients at high risk of poor health outcomes with telehealth during the COVID-19 pandemic: a qualitative study. *BMC Fam Pract* 2021 Apr 08;22(1):69 [FREE Full text] [doi: [10.1186/s12875-021-01408-w](https://doi.org/10.1186/s12875-021-01408-w)] [Medline: [33832422](https://pubmed.ncbi.nlm.nih.gov/33832422/)]
71. Ervin K, Weller-Newton J, Phillips J. Primary healthcare clinicians. *Aust J Prim Health* 2021 Mar 03:158-162. [doi: [10.1071/PY20182](https://doi.org/10.1071/PY20182)] [Medline: [33653506](https://pubmed.ncbi.nlm.nih.gov/33653506/)]
72. Imlach F, McKinlay E, Middleton L, Kennedy J, Pledger M, Russell L, et al. Telehealth consultations in general practice during a pandemic lockdown: survey and interviews on patient experiences and preferences. *BMC Fam Pract* 2020 Dec 13;21(1):269 [FREE Full text] [doi: [10.1186/s12875-020-01336-1](https://doi.org/10.1186/s12875-020-01336-1)] [Medline: [33308161](https://pubmed.ncbi.nlm.nih.gov/33308161/)]
73. Zandi D. Digital health and COVID-19. *Bull World Health Organ* 2020 Nov 01;98(11):731-732 [FREE Full text] [doi: [10.2471/BLT.20.021120](https://doi.org/10.2471/BLT.20.021120)] [Medline: [33177768](https://pubmed.ncbi.nlm.nih.gov/33177768/)]
74. Fatehi F, Wootton R. Telemedicine, telehealth or e-health? A bibliometric analysis of the trends in the use of these terms. *J Telemed Telecare* 2012 Dec;18(8):460-464. [doi: [10.1258/jtt.2012.gth108](https://doi.org/10.1258/jtt.2012.gth108)] [Medline: [23209265](https://pubmed.ncbi.nlm.nih.gov/23209265/)]
75. Recommendations on digital interventions for health system strengthening. World Health Organization. 2019. URL: <https://www.who.int/publications/i/item/9789241550505> [accessed 2021-11-20]
76. Maia MR, Castela E, Pires A, Lapão LV. How to develop a sustainable telemedicine service? A Pediatric Telecardiology Service 20 years on - an exploratory study. *BMC Health Serv Res* 2019 Sep 23;19(1):681 [FREE Full text] [doi: [10.1186/s12913-019-4511-5](https://doi.org/10.1186/s12913-019-4511-5)] [Medline: [31547824](https://pubmed.ncbi.nlm.nih.gov/31547824/)]
77. Atherton H, Brant H, Ziebland S, Bikker A, Campbell J, Gibson A, et al. The potential of alternatives to face-to-face consultation in general practice, and the impact on different patient groups: a mixed-methods case study. *Health Serv Deliv Res* 2018 Jun;6(20):1-200. [doi: [10.3310/hsdr06200](https://doi.org/10.3310/hsdr06200)] [Medline: [29889485](https://pubmed.ncbi.nlm.nih.gov/29889485/)]
78. Thiyagarajan A, Grant C, Griffiths F, Atherton H. Exploring patients' and clinicians' experiences of video consultations in primary care: a systematic scoping review. *BJGP Open* 2020;4(1):bjgpopen20X101020 [FREE Full text] [doi: [10.3399/bjgpopen20X101020](https://doi.org/10.3399/bjgpopen20X101020)] [Medline: [32184212](https://pubmed.ncbi.nlm.nih.gov/32184212/)]
79. Huibers L, Moth G, Carlsen AH, Christensen MB, Vedsted P. Telephone triage by GPs in out-of-hours primary care in Denmark: a prospective observational study of efficiency and relevance. *Br J Gen Pract* 2016 Sep;66(650):e667-e673 [FREE Full text] [doi: [10.3399/bjgp16X686545](https://doi.org/10.3399/bjgp16X686545)] [Medline: [27432608](https://pubmed.ncbi.nlm.nih.gov/27432608/)]
80. Uscher-Pines L, Jones M, Sousa J, Predmore Z, Ober A. The doctor will call me maybe: the uncertain future of audio-only visits and why we need them to address disparities. *The Rand Blog*. 2021 Mar 12. URL: <https://www.rand.org/blog/2021/03/the-doctor-will-call-me-maybe-the-uncertain-future.html> [accessed 2021-11-01]
81. Donaghy E, Atherton H, Hammersley V, McNeilly H, Bikker A, Robbins L, et al. Acceptability, benefits, and challenges of video consulting: a qualitative study in primary care. *Br J Gen Pract* 2019 Sep;69(686):e586-e594 [FREE Full text] [doi: [10.3399/bjgp19X704141](https://doi.org/10.3399/bjgp19X704141)] [Medline: [31160368](https://pubmed.ncbi.nlm.nih.gov/31160368/)]
82. Randhawa RS, Chandan JS, Thomas T, Singh S. An exploration of the attitudes and views of general practitioners on the use of video consultations in a primary healthcare setting: a qualitative pilot study. *Prim Health Care Res Dev* 2019 Jan;20:e5 [FREE Full text] [doi: [10.1017/S1463423618000361](https://doi.org/10.1017/S1463423618000361)] [Medline: [29909798](https://pubmed.ncbi.nlm.nih.gov/29909798/)]
83. Lawson E, Kumar S. The Wass report: moving forward 3 years on. *Br J Gen Pract* 2020 Apr;70(693):164-165 [FREE Full text] [doi: [10.3399/bjgp20X708953](https://doi.org/10.3399/bjgp20X708953)] [Medline: [32217573](https://pubmed.ncbi.nlm.nih.gov/32217573/)]
84. Greenhalgh T, Wherton J, Shaw S, Morrison C. Video consultations for covid-19. *BMJ* 2020 Mar 12;368:m998. [doi: [10.1136/bmj.m998](https://doi.org/10.1136/bmj.m998)] [Medline: [32165352](https://pubmed.ncbi.nlm.nih.gov/32165352/)]
85. Greenhalgh T, Wherton J, Papoutsi C, Lynch J, Hughes G, A'Court C, et al. Beyond adoption: a new framework for theorizing and evaluating nonadoption, abandonment, and challenges to the scale-up, spread, and sustainability of health and care technologies. *J Med Internet Res* 2017 Nov 01;19(11):e367 [FREE Full text] [doi: [10.2196/jmir.8775](https://doi.org/10.2196/jmir.8775)] [Medline: [29092808](https://pubmed.ncbi.nlm.nih.gov/29092808/)]
86. Farr M, Banks J, Edwards HB, Northstone K, Bernard E, Salisbury C, et al. Implementing online consultations in primary care: a mixed-method evaluation extending normalisation process theory through service co-production. *BMJ Open* 2018 Mar 19;8(3):e019966 [FREE Full text] [doi: [10.1136/bmjopen-2017-019966](https://doi.org/10.1136/bmjopen-2017-019966)] [Medline: [29555817](https://pubmed.ncbi.nlm.nih.gov/29555817/)]

87. Breton M, Deville-Stoetzel N, Gaboury I, Smithman MA, Kaczorowski J, Lussier M, et al. Telehealth in primary healthcare: a portrait of its rapid implementation during the COVID-19 pandemic. *Healthc Policy* 2021 Aug;17(1):73-90 [FREE Full text] [doi: [10.12927/hcpol.2021.26576](https://doi.org/10.12927/hcpol.2021.26576)] [Medline: [34543178](https://pubmed.ncbi.nlm.nih.gov/34543178/)]
88. Coronavirus (COVID-19): Guide de soutien pour la mise en œuvre des téléconsultations dans les établissements de santé du Québec dans le contexte de pandémie. Telesante Quebec. 2020. URL: https://telesantequebec.ca/wp-content/uploads/2020/12/20_210_133W_Guide_mise_en_oeuvre_teleconsultations_pandemie_MSSS_V2_VF.pdf [accessed 2021-11-20]
89. Lim J, Broughan J, Crowley D, O'Kelly B, Fawsitt R, Burke MC, et al. COVID-19's impact on primary care and related mitigation strategies: a scoping review. *Eur J Gen Pract* 2021 Dec;27(1):166-175 [FREE Full text] [doi: [10.1080/13814788.2021.1946681](https://doi.org/10.1080/13814788.2021.1946681)] [Medline: [34282695](https://pubmed.ncbi.nlm.nih.gov/34282695/)]
90. Neves AL, Li E, Gupta PP, Fontana G, Darzi A. Virtual primary care in high-income countries during the COVID-19 pandemic: policy responses and lessons for the future. *Eur J Gen Pract* 2021 Dec;27(1):241-247 [FREE Full text] [doi: [10.1080/13814788.2021.1965120](https://doi.org/10.1080/13814788.2021.1965120)] [Medline: [34431426](https://pubmed.ncbi.nlm.nih.gov/34431426/)]
91. Jnr BA. Use of telemedicine and virtual care for remote treatment in response to COVID-19 pandemic. *J Med Syst* 2020 Jun 15;44(7):132 [FREE Full text] [doi: [10.1007/s10916-020-01596-5](https://doi.org/10.1007/s10916-020-01596-5)] [Medline: [32542571](https://pubmed.ncbi.nlm.nih.gov/32542571/)]
92. Car J, Koh GC, Foong PS, Wang CJ. Video consultations in primary and specialist care during the covid-19 pandemic and beyond. *BMJ* 2020 Oct 20;371:m3945. [doi: [10.1136/bmj.m3945](https://doi.org/10.1136/bmj.m3945)] [Medline: [33082127](https://pubmed.ncbi.nlm.nih.gov/33082127/)]
93. Jnr BA. Implications of telehealth and digital care solutions during COVID-19 pandemic: a qualitative literature review. *Inform Health Soc Care* 2021 Mar 02;46(1):68-83. [doi: [10.1080/17538157.2020.1839467](https://doi.org/10.1080/17538157.2020.1839467)] [Medline: [33251894](https://pubmed.ncbi.nlm.nih.gov/33251894/)]
94. Telehealth to expand access to essential health services during the COVID-19 pandemic. Centers for Disease Control and Prevention. 2020 Jun 10. URL: <https://www.cdc.gov/coronavirus/2019-ncov/hcp/telehealth.html> [accessed 2021-12-30]
95. Gerke S, Stern AD, Minssen T. Germany's digital health reforms in the COVID-19 era: lessons and opportunities for other countries. *NPJ Digit Med* 2020;3:94. [doi: [10.1038/s41746-020-0306-7](https://doi.org/10.1038/s41746-020-0306-7)] [Medline: [32685700](https://pubmed.ncbi.nlm.nih.gov/32685700/)]
96. Seto E, Smith D, Jacques M, Morita PP. Opportunities and challenges of telehealth in remote communities: case study of the Yukon telehealth system. *JMIR Med Inform* 2019 Nov 01;7(4):e11353 [FREE Full text] [doi: [10.2196/11353](https://doi.org/10.2196/11353)] [Medline: [31682581](https://pubmed.ncbi.nlm.nih.gov/31682581/)]
97. Egerton T, Nelligan R, Setchell J, Atkins L, Bennell KL. General practitioners' perspectives on a proposed new model of service delivery for primary care management of knee osteoarthritis: a qualitative study. *BMC Fam Pract* 2017 Sep 07;18(1):85 [FREE Full text] [doi: [10.1186/s12875-017-0656-7](https://doi.org/10.1186/s12875-017-0656-7)] [Medline: [28882108](https://pubmed.ncbi.nlm.nih.gov/28882108/)]
98. Eccles A, Hopper M, Turk A, Atherton H. Patient use of an online triage platform: a mixed-methods retrospective exploration in UK primary care. *Br J Gen Pract* 2019 May;69(682):e336-e344 [FREE Full text] [doi: [10.3399/bjgp19X702197](https://doi.org/10.3399/bjgp19X702197)] [Medline: [30910874](https://pubmed.ncbi.nlm.nih.gov/30910874/)]
99. Hammersley V, Donaghy E, Parker R, McNeilly H, Atherton H, Bikker A, et al. Comparing the content and quality of video, telephone, and face-to-face consultations: a non-randomised, quasi-experimental, exploratory study in UK primary care. *Br J Gen Pract* 2019 Sep;69(686):e595-e604 [FREE Full text] [doi: [10.3399/bjgp19X704573](https://doi.org/10.3399/bjgp19X704573)] [Medline: [31262846](https://pubmed.ncbi.nlm.nih.gov/31262846/)]
100. Banks J, Farr M, Salisbury C, Bernard E, Northstone K, Edwards H, et al. Use of an electronic consultation system in primary care: a qualitative interview study. *Br J Gen Pract* 2018 Jan;68(666):e1-e8 [FREE Full text] [doi: [10.3399/bjgp17X693509](https://doi.org/10.3399/bjgp17X693509)] [Medline: [29109115](https://pubmed.ncbi.nlm.nih.gov/29109115/)]
101. Ball SL, Newbould J, Corbett J, Exley J, Pitchforth E, Roland M. Qualitative study of patient views on a 'telephone-first' approach in general practice in England: speaking to the GP by telephone before making face-to-face appointments. *BMJ Open* 2018 Dec 31;8(12):e026197 [FREE Full text] [doi: [10.1136/bmjopen-2018-026197](https://doi.org/10.1136/bmjopen-2018-026197)] [Medline: [30598491](https://pubmed.ncbi.nlm.nih.gov/30598491/)]
102. Hanna L, Fairhurst K. Using information and communication technologies to consult with patients in Victorian primary care: the views of general practitioners. *Aust J Prim Health* 2013;19(2):166-170. [doi: [10.1071/PY11153](https://doi.org/10.1071/PY11153)] [Medline: [22950961](https://pubmed.ncbi.nlm.nih.gov/22950961/)]
103. Newbould J, Exley J, Ball S, Corbett J, Pitchforth E, Roland M. GPs' and practice staff's views of a telephone first approach to demand management: a qualitative study in primary care. *Br J Gen Pract* 2019 May;69(682):e321-e328 [FREE Full text] [doi: [10.3399/bjgp19X702401](https://doi.org/10.3399/bjgp19X702401)] [Medline: [31015225](https://pubmed.ncbi.nlm.nih.gov/31015225/)]
104. The potential of frequently used information technologies during the pandemic. Pan American Health Organization, Institutional Repository for Information Sharing. URL: <https://iris.paho.org/handle/10665.2/52023> [accessed 2021-11-20]
105. Atenção primária à saúde. Organização Pan-Americana da Saúde. URL: <https://www.paho.org/pt/topicos/atencao-primaria-saude> [accessed 2021-11-20]
106. Programme in Health Services and Systems Research. Duke NUS Medical School. URL: <https://www.duke-nus.edu.sg/hssr> [accessed 2021-11-20]
107. Chong JL, Lim KK, Matchar DB. Population segmentation based on healthcare needs: a systematic review. *Syst Rev* 2019 Aug 13;8(1):202 [FREE Full text] [doi: [10.1186/s13643-019-1105-6](https://doi.org/10.1186/s13643-019-1105-6)] [Medline: [31409423](https://pubmed.ncbi.nlm.nih.gov/31409423/)]
108. Lillrank P, Groop P, Malmström TJ. Demand and supply-based operating modes--a framework for analyzing health care service production. *Milbank Q* 2010 Dec;88(4):595-615 [FREE Full text] [doi: [10.1111/j.1468-0009.2010.00613.x](https://doi.org/10.1111/j.1468-0009.2010.00613.x)] [Medline: [21166870](https://pubmed.ncbi.nlm.nih.gov/21166870/)]
109. Salvador PTCDO, Alves KYA, Costa TDD, Lopes RH, Oliveira LVE, Rodrigues CCFM. Contribuições da scoping review na produção da área da saúde: reflexões e perspectivas. *REDCPS* 2021;6:1-8. [doi: [10.5935/2446-5682.20210058](https://doi.org/10.5935/2446-5682.20210058)]

Abbreviations

ICT: information and communication technology

JBI: Joanna Briggs Institute

PCC: Population, Concept, and Context

PHC: primary health care

PRISMA: Preferred Reporting Items for Systematic Reviews and Meta-analyses

PRISMA-ScR: Preferred Reporting Items for Systematic Reviews and Meta-analyses Extension for Scoping Reviews

WHO: World Health Organization

Edited by A Kushniruk; submitted 02.12.21; peer-reviewed by E van der Velde, B Wolfe-Piggott, M Peyroteo; comments to author 23.12.21; revised version received 11.01.22; accepted 21.03.22; published 31.05.22.

Please cite as:

Silva CRDV, Lopes RH, de Goes Bay Jr O, Martiniano CS, Fuentealba-Torres M, Arcêncio RA, Lapão LV, Dias S, Uchoa SADC

Digital Health Opportunities to Improve Primary Health Care in the Context of COVID-19: Scoping Review

JMIR Hum Factors 2022;9(2):e35380

URL: <https://humanfactors.jmir.org/2022/2/e35380>

doi: [10.2196/35380](https://doi.org/10.2196/35380)

PMID: [35319466](https://pubmed.ncbi.nlm.nih.gov/35319466/)

©Cícera Renata Diniz Vieira Silva, Rayssa Horácio Lopes, Osvaldo de Goes Bay Jr, Claudia Santos Martiniano, Miguel Fuentealba-Torres, Ricardo Alexandre Arcêncio, Luís Velez Lapão, Sonia Dias, Severina Alice da Costa Uchoa. Originally published in JMIR Human Factors (<https://humanfactors.jmir.org>), 31.05.2022. This is an open-access article distributed under the terms of the Creative Commons Attribution License (<https://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Human Factors, is properly cited. The complete bibliographic information, a link to the original publication on <https://humanfactors.jmir.org>, as well as this copyright and license information must be included.

Review

The Impact of Perioperative Remote Patient Monitoring on Clinical Staff Workflows: Scoping Review

Maria Alejandra León¹, MSc; Valeria Pannunzio¹, MSc; Maaïke Kleinsmann¹, PhD

Department of Design, Organization and Strategy, Faculty of Industrial Design Engineering, Delft University of Technology, Delft, Netherlands

Corresponding Author:

Valeria Pannunzio, MSc

Department of Design, Organization and Strategy

Faculty of Industrial Design Engineering

Delft University of Technology

Landbergstraat 15, 2628 CE Delft

Delft,

Netherlands

Phone: 31 152781460

Email: v.pannunzio@tudelft.nl

Abstract

Background: Remote patient monitoring (RPM) interventions are being increasingly implemented in health care environments, given their benefits for different stakeholders. However, the effects of these interventions on the workflow of clinical staff are not always considered in RPM research and practice.

Objective: This review explored how contemporary RPM interventions affect clinical staff and their workflows in perioperative settings.

Methods: We conducted a scoping review of recent articles reporting the impact of RPM interventions implemented in perioperative settings on clinical staff and their workflow. The databases accessed were Embase and PubMed. A qualitative analysis was performed to identify the main problems and advantages that RPM brings to staff, in addition to the approaches taken to evaluate the impact of those interventions. Different themes were identified in terms of the challenges of RPM for clinical staff as well as in terms of benefits, risk-reduction strategies, and methods for measuring the impact of these interventions on the workflow of clinical staff.

Results: A total of 1063 papers were found during the initial search, of which 21 (1.98%) met the inclusion criteria. Of the 21 included papers, 15 (71%) focused on evaluating new RPM systems, 4 (19%) focused on existing systems, and 2 (10%) were reviews.

Conclusions: The reviewed literature shows that the impact on staff work experience is a crucial factor to consider when developing and implementing RPM interventions in perioperative settings. However, we noticed both underdevelopment and lack of standardization in the methods for assessing the impact of these interventions on clinical staff and their workflow. On the basis of the reviewed literature, we recommend the development of more robust methods for evaluating the impact of RPM interventions on staff experience in perioperative care; the adoption of a stronger focus on transition management when introducing these interventions in clinical practice; and the inclusion of longer periods of assessment, including the evaluation of long-term goals.

(*JMIR Hum Factors* 2022;9(2):e37204) doi:[10.2196/37204](https://doi.org/10.2196/37204)

KEYWORDS

remote patient monitoring; telemonitoring; workflow; nurses; physicians; perioperative care; perioperative medicine; telehealth; mobile phone

Introduction

Background

Remote patient monitoring (RPM) interventions allow patients to be continuously monitored at a distance and beyond the physical borders of the hospital or health care institution [1]. RPM interventions have been used to monitor patients within clinical settings (eg, in intensive care environments) or outside of care facilities (eg, in the patients' homes). Moreover, RPM has been used for delivering care for multiple health conditions, from heart failure [2] to diabetes [3] and skin problems [4].

RPM interventions can provide 24-hour care as they can collect data continuously and alert specialists when certain parameters are outside the standard thresholds [5]. This can enable real-time adjustments, timely decisions, and improved care. RPM as a field has also enjoyed an unprecedented acceleration as a consequence of the COVID-19 pandemic, which has stimulated the adoption of remote care to minimize face-to-face interactions between patients and staff [6]. In the perioperative setting, RPM can be useful for assessing physical conditions preoperatively or monitoring patients' recovery after discharge. Although RPM applications in this domain are still relatively novel, encouraging results are driving an increased interest from researchers and practitioners.

An example of the application of RPM technologies to perioperative care was offered by Atilgan et al [7], who evaluated a system comprising monitoring devices collecting several vital signs (including blood pressure, heart rate, oxygen saturation, body temperature, blood glucose, and electrocardiography) and a mobile app providing medication reminders, suggested daily life activities, diet and nutrition plans, and web-based visit capabilities. Vital parameters were measured in patients who had undergone cardiac surgery after discharge and automatically transferred to a telemedicine team for assessment. Overall, the authors reported the RPM intervention to have resulted in high patient satisfaction, prevention of incorrect medications and dosages, prevention of rehospitalization, and early detection of potentially life-threatening complications.

Much of the available research on RPM interventions in the perioperative domain focuses on the effects of RPM on patients [8-11] and describes its advantages, especially in terms of clinical outcomes and efficiency gains [12-14]. Some studies have also addressed the benefits for health care providers, such as hospitals, nursing homes, and other entities. These studies tend to focus on the economic benefits for providers, for instance through reductions in hospitalizations and thus, in the use of resources [15,16]. However, there is limited knowledge of the benefits and limitations of RPM for clinical staff.

Objectives

This research seeks to evaluate the impact of RPM interventions on the workflow of clinical staff in the context of perioperative care. To explain what we mean by *workflow*, we follow Carayon et al [17], who defined workflow as “the flow of people, equipment, information, and tasks, in different places, at different levels, at different timescales continuously and

discontinuously, that are used or required to support the goals of the work domain.” This means that we aimed to evaluate the impact of RPM-related tasks in combination with previously existing activities. In this paper, the words *clinical staff* will be used when referring to both nurses and specialists. To investigate the impact of RPM on the workflow of clinical staff, a human factor perspective was adopted in this review. As mentioned by Hignett et al [18], human factors help in understanding the interactions between humans and the elements of a system to optimize its performance and human well-being.

This scoping review sought to answer the following overall research question: What is the impact of perioperative RPM interventions on the workflow of clinical staff? To answer this main question, we developed the following subresearch questions: (1) What are the problems and challenges of perioperative RPM interventions for clinical staff from a workflow perspective? (2) What are the benefits of perioperative RPM interventions for clinical staff from a workflow perspective? (3) What strategies are implemented or proposed to overcome the problems that perioperative RPM interventions present to the workflow of clinical staff? (4) How is the impact of perioperative RPM interventions on the workflow of clinical staff evaluated and measured?

Methods

Overview

This scoping review followed the PRISMA-ScR (Preferred Reporting Items for Systematic Reviews and Meta-Analyses extension for Scoping Reviews) checklist [19]. As the review focuses on collecting and comparing workflow-related insights from recent RPM literature rather than on drawing conclusions on specific outcomes, the risk of bias in the results of the included studies was not assessed. Conversely, the risk of bias in the synthesis of the literature review findings was considered. Specifically, the risk of bias owing to missing results was assessed by MAL and VP through the framework for assessing the risk of bias owing to missing results in a synthesis offered in the Cochrane Handbook for Systematic Reviews of Interventions [20]. The results of this assessment are discussed in the *Limitations* section.

Selection Criteria and Search Strategy

The databases used were PubMed and Embase. To define the inclusion criteria, key concepts were selected. For each of them, keywords were defined to guide the search strategy (Textbox 1). For the keywords of each concept, the logical operator *OR* was included to consider all the possibilities, whereas the logical operator *AND* was used between concepts. The full queries in both databases are presented in Multimedia Appendix 1. Finally, the search included articles that were written in English between January 2015 and March 2021. This was chosen to obtain a picture of contemporary RPM interventions as this review focuses on current challenges and opportunities. The search was conducted during the last week of March 2021.

The articles resulting from this search were screened based on the following inclusion criteria: (1) the inclusion of RPM

interventions for perioperative care and (2) the mention of the impact on the workflow of clinical staff.

The criteria were used for 2 iterations of screening: the first was based on the title and abstract of the articles, and the second was based on the full text.

Textbox 1. Concepts included in the literature search.

- A keyword can have some variations (plural or singular form or simple or continuous verb form). An asterisk (*) is used for the search algorithm in the database to find all possible variations of a certain word.
 - Remote patient monitoring: remote monitor*; telemedicine; telemonitoring; telehealth, remote follow-up; eHealth; remote consultation; remote sensing technology; self-monitor*
 - Workflow: workflow; outcome and process assessment, health care; task performance and analysis; workflow; staffing; attitude of health personnel; alarm fatigue*; alert fatigue; professional burnout, workload; patient care management; nursing process*; clinical competence; caregiver burden; time and motion studies; work simplification; practice patterns, nurses; nursing audit
 - Perioperative care: surgical procedures, operative; general surgery; perioperative; surgery; post-operative; post-discharge

Review Process and Analysis

Our main categories were established (Textbox 2) to analyze the studies, namely challenges and problems, benefits, risk-reduction strategies, and evaluation methods. These were based on the main goals of this research and the research questions.

The articles were reviewed by MAL, who was also responsible for data extraction. Subsequently, the first step of the analysis was performed by classifying the results into the chosen categories. The second step consisted of creating different themes per category. This step required several iterations to obtain the final set of themes.

Textbox 2. Categories used for data extraction.

- Problems and challenges of remote patient monitoring (RPM) interventions for clinical staff: includes the problems shown regarding RPM interventions for clinical staff.
- Benefits of RPM interventions for clinical staff: includes the benefits concerning RPM interventions for clinical staff.
- Risk-reduction strategies regarding RPM interventions for clinical staff: includes solutions tested to tackle some of the problems brought by the introduction of RPM interventions and some of the proposals suggested.
- Methods to measure and quantify the impact of RPM interventions on clinical staff: includes the methods used to determine the impact of RPM interventions on clinical staff's tasks and workflow. It entails the variables and measures collected and analyzed.

Results

Overview

A total of 1063 articles were identified after searching both databases, of which 1007 (94.73%) were left after deduplication. Of these 1007 articles, 137 (13.6%) fulfilled the first round of selection, and 21 (2.09%) passed the final round of selection (Figure 1).

In general, the articles included in this review were experimental or observational studies. Of the 21 articles, 15 (71%) involved the evaluation of a design intervention (an RPM model, tool, or service), 4 (19%) consisted of an analysis of already implemented interventions, and the remaining 2 (10%) were reviews. The references and articles analyzed in these 2 reviews did not include any of the other selected articles in this scoping review.

The studies focused on a wide range of patient cohorts and surgical specialties, including orthopedic, bariatric, and oncological surgery. Most of these studies (20/21, 95%) focused on adult patients (aged >18 years). The described RPM interventions ranged from 1 to 45 months of duration.

In addition, the articles presented different types of RPM interventions, ranging from e-tools used only by the clinical staff to services and models that incorporated devices and platforms for both patients and specialists. Moreover, most of the interventions contemplated nursing staff as the main actors responsible for remote care and included physicians for specific tasks or just in case a more detailed and in-depth analysis of the patient's data was needed.

Not all the included studies contained information on all the categories established. For example, the included reviews hardly mentioned the methodologies used to assess the impact of different RPM interventions on the workflow of the clinical staff.

Once the data were extracted from the articles, they were classified into the 4 categories. To better understand each category, different themes were defined (Figure 2) based on the similarity of the topics addressed in each of the articles. Figure 2 presents an overview of this classification, where each category is labeled with a different color. By means of a gradient in the color's intensity, it is possible to show the quantity of papers that touch on each of the proposed themes. In this case, more saturated colors represent more papers mentioning information relevant to a specific theme. The results for each category are discussed in detail in the following sections.

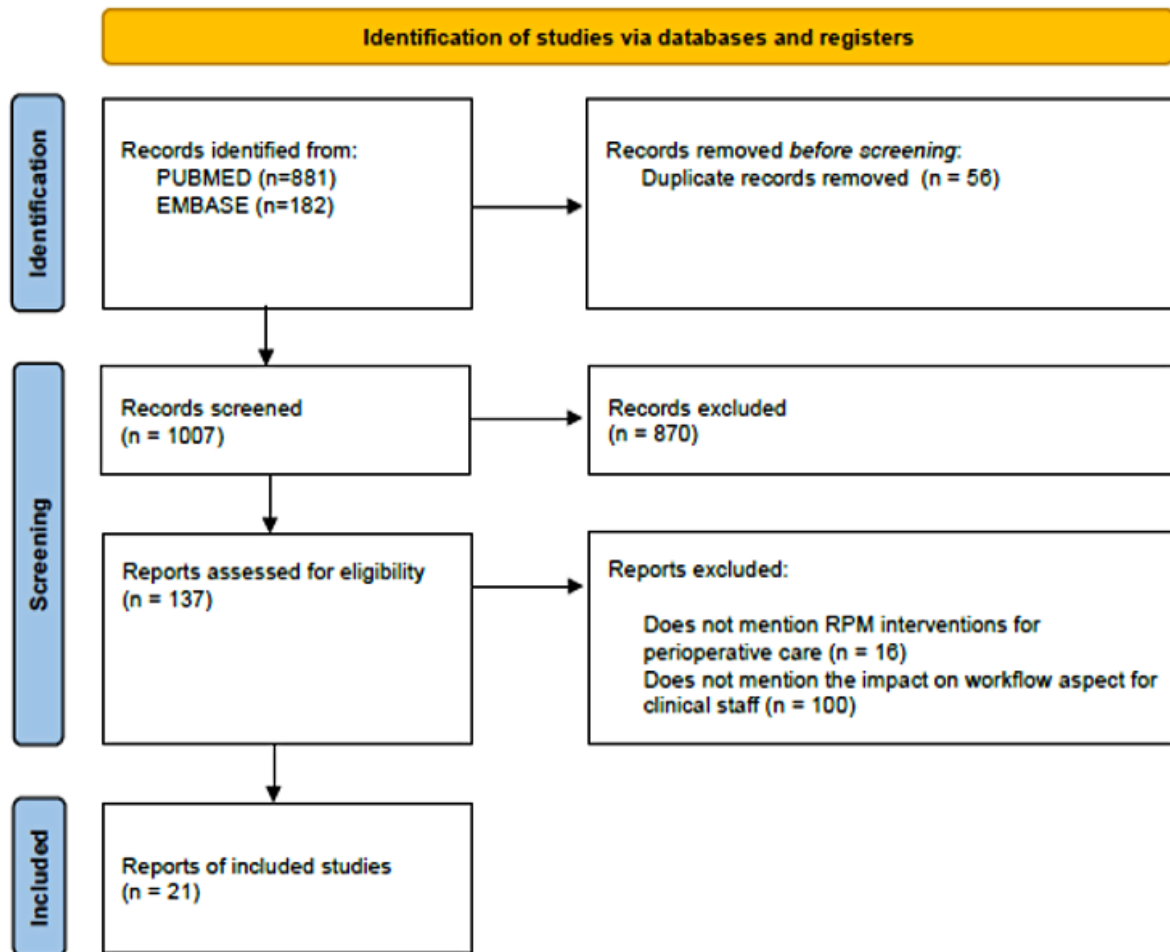
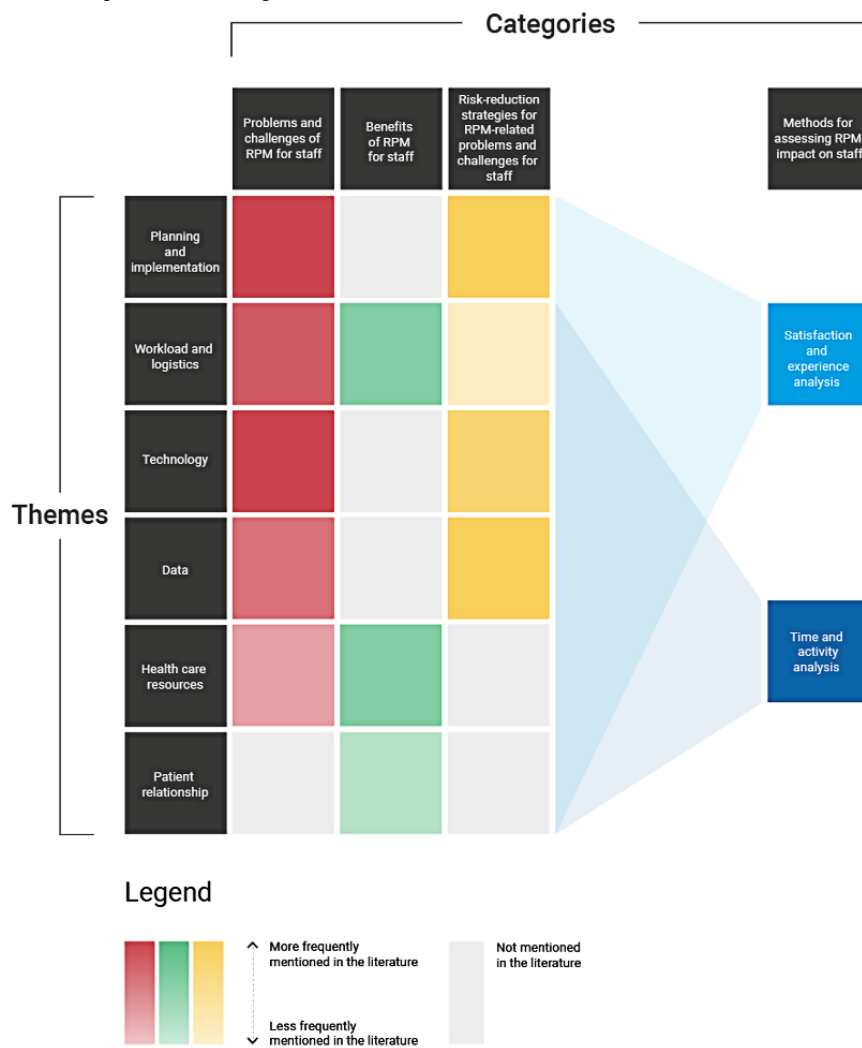
Figure 1. Flowchart of the scoping review process and the inclusion and exclusion criteria. RPM: remote patient monitoring.

Figure 2. Heat map of the review results organized by categories (each corresponding to a research question) and themes (recurring topics touched on in the included studies). RPM: remote patient monitoring.



Category 1: Problems and Challenges of RPM for Clinical Staff

On the basis of the articles analyzed, 5 main themes regarding RPM challenges from the viewpoint of clinical staff were identified (Table 1). The first theme was planning and implementation. Planning is a complex task in health care given the diversity of the stakeholders involved and their needs. RPM projects do not always involve or consider the complex context in which these interventions have to be implemented. This often leads to ambiguity in tasks and roles and, thus, to lack of clarity and structure in the workflow of the clinical staff.

The second theme was workload and logistics. Some staff members do not feel comfortable with the new *behind-the-desk* activities, which can result in unpredictable and emergent tasks when RPM systems register values outside the thresholds. Moreover, data analysis may require more than one specialist, making the workflow more complex. In addition, RPM is

perceived as bringing more work, which adds to the existing schedule.

The third theme was technology. Systems might not be user-friendly, and different technical malfunctions may arise, which may require extra expertise from clinical staff.

The fourth theme was data, which can produce more informed decisions but also increase time and be burdensome to analyze. Moreover, it can be hard to keep all the data under 1 platform, so the staff may need to analyze multiple fragments of information to provide remote care.

The last theme was health care resources, intended as the new resources that RPM interventions require. Moreover, the aforementioned ambiguity in tasks determines a lack of clarity regarding reimbursement policies.

A detailed overview of the reported challenges for each category is provided in Table 1.

Table 1. Overview of problems and challenges of remote patient monitoring (RPM) interventions for clinical staff.

Theme and description	Studies
Planning and implementation	
<ul style="list-style-type: none"> Lack of previous user testing 	<ul style="list-style-type: none"> Harsha et al [21]
<ul style="list-style-type: none"> Lack of planning or inadequate planning <ul style="list-style-type: none"> Lack of contemplation of changes in workflow (tasks, competences, responsibilities, and roles) Emergence of unanticipated tasks No standardization in practices and no clear guidelines Noncompatibility with current practices No clear definition of time for tasks No long-term care coordination Services are implemented before all the resources are available and prepared 	<ul style="list-style-type: none"> Das et al [22] Davoody and Hägglund [23] Harsha et al [21] Ke et al [24] Leppla et al [25] Sanger et al [26] Timmerman et al [27] Wiadji et al [28]
<ul style="list-style-type: none"> Lack of resource analysis (“readiness level”) <ul style="list-style-type: none"> No clear overview of required skills No consideration of staff experience No clarity on resource accessibility (whether clinical staff is adequately equipped) 	<ul style="list-style-type: none"> Ke et al [24] Parkes et al [29] Rothgangel et al [30] Wiadji et al [28]
<ul style="list-style-type: none"> Lack of multidisciplinary awareness <ul style="list-style-type: none"> Uncontemplated users, nonusers, and other actors affected Limited or poor communication and coordination among users Poor task planning (tasks overlapping and no consideration for the need of staff to attend to 1 patient at a time) Disregard for the specificities of different specialties and wards (eg, cardiovascular and pediatric) 	<ul style="list-style-type: none"> Harsha et al [21] Leppla et al [25] Makhni et al [31] Parkes et al [29] Wiadji et al [28]
<ul style="list-style-type: none"> Lack of compliance and engagement <ul style="list-style-type: none"> Lack of involvement of stakeholders in planning Fear of conflict of interest Lack of promotion and motivation among staff Decrease of use of systems over time Resistance to change Specialists and rural hospitals, among others, feeling threatened to be replaced 	<ul style="list-style-type: none"> Downey et al [32] Harsha et al [21] McMullen et al [33] Parkes et al [29] Rothgangel et al [30] Sharif et al [34] Timmerman et al [27] Wiadji et al [28]
Workload and logistics	
<ul style="list-style-type: none"> High workload <ul style="list-style-type: none"> New tasks as an addition and not a replacement Telehealth tasks are perceived to be labor-intensive (“More administrative work in arranging telehealth than meets the eyes”) Tracking patients takes too much time (because of subtasks such as setting up appointments, billing, mailing, analyzing, reviewing transmissions, documenting in the EMR^a, and physician contact) Remote patients are not considered as part of “normal flow” (ignored for workload calculation) Potentially adding an unnecessary step when patient attention is needed (immediate patient check by GP^b instead of data follow-up by nurse) Documentation is burdensome 	<ul style="list-style-type: none"> Brophy [35] Das et al [22] Dunphy et al [36] Harsha et al [21] Ke et al [24] Leppla et al [25] Makhni et al [31] McMullen et al [33] Parkes et al [29] Sharif et al [34] Wiadji et al [28]
<ul style="list-style-type: none"> Disruption in workflow <ul style="list-style-type: none"> Unpredictable, emergent tasks High memory load Mistakes on interrupted activities Unanswered or unplanned calls 	<ul style="list-style-type: none"> Das et al [22] Downey et al [32] Harsha et al [21] Sanger et al [26]
<ul style="list-style-type: none"> Nonurgent tasks emerge outside working hours 	<ul style="list-style-type: none"> Ke et al [24]
<ul style="list-style-type: none"> Need of trustworthy professionals for data analysis <ul style="list-style-type: none"> Nurses sometimes need to consult with physicians 	<ul style="list-style-type: none"> Leppla et al [25]
<ul style="list-style-type: none"> Fear of infringing on other providers’ patient care 	<ul style="list-style-type: none"> Brophy [35]

Theme and description	Studies
<ul style="list-style-type: none"> Stress because of pressure for timely responses to multiple issues 	<ul style="list-style-type: none"> Das et al [22] McMullen et al [33] Parkes et al [29]
Technology	
<ul style="list-style-type: none"> Difficulties in use of e-tools <ul style="list-style-type: none"> Not user-friendly No experience or training Technical problems <ul style="list-style-type: none"> Troubleshooting and malfunctions Connection issues (eg, congestion, no signal, and delays) Not compatible with current software Deficient communication <ul style="list-style-type: none"> Inappropriate means of communication Hard to establish “personal connection” for communicating bad news or managing conflict with patients New medical-legal situations (patients might misunderstand information or take it out of context) RPM interventions might not be suitable to all the patients RPM does not offer monitoring to the same extent as in-hospital monitoring <ul style="list-style-type: none"> No physical examination Cannot assess if patient does self-monitoring or prescribe activities correctly 	<ul style="list-style-type: none"> Brophy [35] Das et al [22] Davoody and Hägglund [23] Parkes et al [29] Rothgangel et al [30] Sousa et al [37] Timmerman et al [27] Augestad et al [38] Brophy [35] Harsha et al [21] Makhni et al [31] Timmerman et al [27] Augestad et al [38] Dunphy et al [36] Ke et al [24] Leppla et al [25] Makhni et al [31] Parkes et al [29] Wiadji et al [28] Dunphy et al [36] Ke et al [24]
Data	
<ul style="list-style-type: none"> False or insignificant alarms or overreaction <ul style="list-style-type: none"> Stress by constant sound Turning devices off or not using them Unclear data and meaning <ul style="list-style-type: none"> Require extensive analysis Overabundance of data No flag data Missing connection among data No clear “holistic” impression of patients <ul style="list-style-type: none"> Lack of data integration with EMR and other existing platforms Not all the reports generated by the system are consulted by physicians Low reliability of patient monitoring <ul style="list-style-type: none"> Incomplete data Incorrect measurements Legal issues (eg, privacy, firewall, and licenses) 	<ul style="list-style-type: none"> Downey et al [32] Harsha et al [21] Richards et al [39] Das et al [22] Leppla et al [25] Sharif et al [34] Das et al [22] Semple et al [40] Sharif et al [34] Timmerman et al [27] Leppla et al [25] Sharif et al [34] Brophy [35] Das et al [22] Ke et al [24] Makhni et al [31] Semple et al [40]
Health care resources	

Theme and description	Studies
<ul style="list-style-type: none"> Lack of funding <ul style="list-style-type: none"> Higher costs than budget Nonsustainable billing rates No clinic income established Higher payment for in-hospital visits 	<ul style="list-style-type: none"> Das et al [22] Brophy [35] Harsha et al [21] Makhni et al [31] Wiadji et al [28]
<ul style="list-style-type: none"> Demand of new or more resources 	<ul style="list-style-type: none"> Das et al [22] Makhni et al [31]
<ul style="list-style-type: none"> Difficult to quantify quality and effort 	<ul style="list-style-type: none"> Wiadji et al [28]
<ul style="list-style-type: none"> Unclear compensation or reimbursement policies <ul style="list-style-type: none"> Telehealth can take up the same amount of time for significantly less remuneration 	<ul style="list-style-type: none"> Brophy [35] Ke et al [24] Semple et al [40]

^aEMR: electronic medical record.

^bGP: general practitioner.

Category 2: Benefits of RPM for Clinical Staff

For the *benefits* category, 3 main themes were identified as relevant (Table 2). The first theme was the improvement that RPM brings regarding workload and logistics as it allows for the definition of guidelines for more consistent care pathways. This also includes improvements in data management and

analysis, which produces timely detection and treatment of conditions.

The second theme was health care resources, which can be operated more effectively with the reduction of in-hospital visits and stays and the possibility of extending coverage of care.

Finally, patient relationship can be improved by increasing satisfaction and convenience of care.

Table 2. Overview of benefits of remote patient monitoring interventions for clinical staff.

Theme and description	Studies
Workload and logistics	
<ul style="list-style-type: none"> Care pathways are standardized <ul style="list-style-type: none"> More systematic and consistent activities 	<ul style="list-style-type: none"> Brophy [35] McMullen et al [33]
<ul style="list-style-type: none"> Reducing the incidence of duplicate documentation 	<ul style="list-style-type: none"> Jansson et al [41]
<ul style="list-style-type: none"> Promote collaboration among health care specialists <ul style="list-style-type: none"> New and appropriate means to hold clinical meetings Patient information can be made accessible to the caregivers involved 	<ul style="list-style-type: none"> Sharif et al [34] Wiadji et al [28]
<ul style="list-style-type: none"> Reduce time to reach a clinical decision <ul style="list-style-type: none"> Shortens face-to-face consultation time Patients are better prepared for the appointment 	<ul style="list-style-type: none"> Ke et al [24] Sharif et al [34]
<ul style="list-style-type: none"> Improve sense-making of data <ul style="list-style-type: none"> Include more sources for analyzing patients' clinical condition (current state, feedback, and patients' experience and feeling) Reassuring system based on predefined algorithms for clinical support and suggestions Increased detection of events Real-time monitoring of symptoms over a prolonged period 	<ul style="list-style-type: none"> Jansson et al [41] Ke et al [24] Leppla et al [25] Makhni et al [31] McMullen et al [33] Parkes et al [29] Richards et al [39] Sharif et al [34] Timmerman et al [27]
Health care resources	
<ul style="list-style-type: none"> Reduce workload 	<ul style="list-style-type: none"> Leppla et al [25] Parkes et al [29]
<ul style="list-style-type: none"> Can reduce costs <ul style="list-style-type: none"> Prevents unnecessary visits and health care use Reduces tests and investigations 	<ul style="list-style-type: none"> Augestad et al [38] Makhni et al [31] Parkes et al [29] Sharif et al [34]
<ul style="list-style-type: none"> Increase accessibility <ul style="list-style-type: none"> More patients can be taken care of More hospitals (eg, rural and remote) can track patients Customizable service (awareness of unique individual challenges) 	<ul style="list-style-type: none"> Augestad et al [38] Brophy [35] Das et al [22] Davoody and Hägglund [23] McMullen et al [33] Timmerman et al [27]
<ul style="list-style-type: none"> Allow for a new form of triage for better assessment of patients and resource allocation 	<ul style="list-style-type: none"> Sharif et al [34] Wiadji et al [28]
Patient relationship	
<ul style="list-style-type: none"> Increase patient satisfaction and convenience 	<ul style="list-style-type: none"> Augestad et al [38] Dunphy et al [36] Parkes et al [29] Sharif et al [34]
<ul style="list-style-type: none"> Increase awareness of patient's daily life 	<ul style="list-style-type: none"> Das et al [22] Davoody and Hägglund [23]

Category 3: Risk-Reduction Strategies Regarding RPM for Clinical Staff

This category is about strategies to overcome and minimize the challenges that RPM interventions bring about to clinical staff (Table 3). For ease of reference, we refer to risk-reduction strategies related to the introduction of RPM interventions as

strategies. First, we listed strategies regarding planning and implementation of RPM interventions. Most of the included studies (14/21, 67%) mentioned the value of involving the relevant stakeholders in these processes to understand their needs and the repercussions of the introduction of the RPM intervention on their workflow. Stakeholders' involvement and participatory approaches were also deemed useful to assess the

resources necessary for RPM interventions, the possible risks associated with them, and the need for possible changes to the implementation plans. Finally, training and establishment of protocols (regarding activities, communication, time, and resources) help in risk reduction during implementation and increase the chances of success and adoption.

Second, we listed strategies regarding workload and logistics. Several included studies (8/21, 38%) suggested the creation of new roles for nurses and teams for the remote care of patients, where specialists would be consulted only in special cases. Some strategies to avoid an increase in workload for nursing

staff included facilitating collaboration between actors and helping them plan their tasks.

The third theme was technology, which should be user-friendly, interoperable with existing devices and systems, and allow for automatic data collection.

Finally, we identified the theme of data. To avoid the analysis of RPM data being burdensome for staff, smart systems based on customizable alerts were proposed to prevent resource overuse and the incidence of false alarms. These should include measurements from different devices or sources and be presented to the relevant staff in an actionable and understandable way to avoid extra time and burden.

Table 3. Overview of risk-reduction strategies regarding remote patient monitoring (RPM) interventions for clinical staff.

Theme and description	Studies
Planning and implementation	
<ul style="list-style-type: none"> Develop an integrated governance structure <ul style="list-style-type: none"> Involve all actors concerned with patient management (co-design and participatory practices) Set clear objectives, success metrics, and methods to measure them 	<ul style="list-style-type: none"> Das et al [22] Harsha et al [21] Ke et al [24] Leppla et al [25] McMullen et al [33] Parkes et al [29] Sanger et al [26] Semple et al [40] Timmerman et al [27]
<ul style="list-style-type: none"> Determine health care resource use in terms of the following: <ul style="list-style-type: none"> Clinical staff and skills Tasks and their timing (to avoid invisible or additional work, time, roles or teams, or an inadequate alert response) Awareness of the multidisciplinary environment Plan for problem solving and changes needed Time for solving technical or general problems Devices and structure 	<ul style="list-style-type: none"> Brophy [35] Das et al [22] Ke et al [24] Leppla et al [25] Parkes et al [29] Richards et al [39] Timmerman et al [27] Wiadji et al [28]
<ul style="list-style-type: none"> Define practice standards, policies, and best practices in terms of the following: <ul style="list-style-type: none"> Workflow Documentation Communication pathways Measurements Types of data collected Impact on the clinical staff's well-being (clinical staff's attitudes, performance, and overall service satisfaction) 	<ul style="list-style-type: none"> Augestad et al [38] Das et al [22] Harsha et al [21] Jansson et al [41] Ke et al [24] Leppla et al [25] Sanger et al [26] Semple et al [40] Timmerman et al [27] Wiadji et al [28]
<ul style="list-style-type: none"> Risk assessment <ul style="list-style-type: none"> Perform adequate device testing Contemplate technical or general problems (extra time) 	<ul style="list-style-type: none"> Brophy [35] Das et al [22] Leppla et al [25] Richards et al [39] Timmerman et al [27]
<ul style="list-style-type: none"> Consider current state and context <ul style="list-style-type: none"> Plan according to resources, program, location, dynamics (within the hospital and among clinical staff), and schedules (consider "less busy" and "very busy" times) Customize interventions for integration with existing clinical dynamics and tools 	<ul style="list-style-type: none"> Brophy [35] Das et al [22] Davoody and Hägglund [23] Jansson et al [41] McMullen et al [33] Richards et al [39] Sousa et al [37]
<ul style="list-style-type: none"> Definition of reimbursement policies <ul style="list-style-type: none"> Automatically track time for standardization Consider financial or nonfinancial options (awards and acknowledgments) Automatically measure time to determine billing Include billing functionalities in the intervention 	<ul style="list-style-type: none"> Das et al [22] Wiadji et al [28]

Theme and description	Studies
<ul style="list-style-type: none"> • Training staff on tools and protocols <ul style="list-style-type: none"> • Promote enthusiasm, value, and importance among medical staff regarding RPM 	<ul style="list-style-type: none"> • Brophy [35] • Das et al [22] • Downey et al [32] • Jansson et al [41] • Leppla et al [25] • Makhni et al [31] • McMullen et al [33] • Rothgangel et al [30] • Semple et al [40] • Sousa et al [37] • Timmerman et al [27] • Wiadji et al [28]
Workload and logistics	
<ul style="list-style-type: none"> • Devise a primary nursing-based model (physicians for emergencies and medical decisions) • Allow for easy collaboration between the different actors • Create dedicated teams for RPM interventions • Include planning tools for routines and tasks <ul style="list-style-type: none"> • Define goals for tasks to make progress clear • Externalize tasks <ul style="list-style-type: none"> • Have specialized centers for data analysis and alarm reviews • Ensure accessibility to patients' contact details (to facilitate appointment scheduling and remote consultations) • Make e-tools available in different languages 	<ul style="list-style-type: none"> • Leppla et al [25] • Davoody and Hägglund [23] • Leppla et al [25] • Leppla et al [25] • Davoody and Hägglund [23] • Leppla et al [25] • Jansson et al [41] • Ke et al [24] • Brophy [35]
Technology	
<ul style="list-style-type: none"> • Provide appropriate support and access to software and technology for both patients and specialists <ul style="list-style-type: none"> • Ensure compatibility with different smartphones and tablets • Ensure QoS^a support • Integrate with current technologies <ul style="list-style-type: none"> • Interoperable and compatible with other or existing devices and systems • Guarantee a seamless connection between RPM platform and staff's EMR^b system • Ensure automatic measurements and documentation • Develop user-friendly tools for clinical staff and patients 	<ul style="list-style-type: none"> • Dunphy et al [36] • Rothgangel et al [30] • Wiadji et al [28] • Harsha et al [21] • Harsha et al [21] • Leppla et al [25] • McMullen et al [33] • Rothgangel et al [30] • Das et al [22] • Ke et al [24] • Sanger et al [26] • Augestad et al [38] • Brophy [35] • Davoody and Hägglund [23] • Leppla et al [25] • McMullen et al [33] • Timmerman et al [27]
Data	

Theme and description	Studies
<ul style="list-style-type: none"> Alert-based follow-up protocol <ul style="list-style-type: none"> Continuous data collection (24-hour data) but data analysis focused on alerts by patient prioritization and event-triggered assessment (identify main events to follow) Automatic event classification and suggestions for corrective actions Providing memory aids to staff for interrupted tasks 	<ul style="list-style-type: none"> Dunphy et al [36] Ke et al [24] McMullen et al [33] Richards et al [39] Sanger et al [26]
<ul style="list-style-type: none"> Customizable data collection <ul style="list-style-type: none"> According to treatment, acuity, goal, progress, and diagnosis (identify high-risk patients to determine extra measures needed) 	<ul style="list-style-type: none"> Das et al [22] Davoody and Hägglund [23] Downey et al [32] Jansson et al [41] Ke et al [24] McMullen et al [33] Rothgangel et al [30]
<ul style="list-style-type: none"> Present easy-to-interpret and actionable data <ul style="list-style-type: none"> Filter data (“noise cancellation” and false positives) Provide comparison of individual scores with “standard values” of comparable patients 	<ul style="list-style-type: none"> Dunphy et al [36] Leppla et al [25] McMullen et al [33] Rothgangel et al [30] Sanger et al [26]
<ul style="list-style-type: none"> Incorporate different kinds of measurements (from different physiological variables) Include historical patients’ data 	<ul style="list-style-type: none"> Davoody and Hägglund [23] Dunphy et al [36] Jansson et al [41] McMullen et al [33] Rothgangel et al [30] Sanger et al [26]
<ul style="list-style-type: none"> More effective use of patients’ data <ul style="list-style-type: none"> Use RPM data to guide future medical appointments Use RPM data to assess eligibility for procedures, possible risks, and outcomes 	<ul style="list-style-type: none"> Dunphy et al [36] Jansson et al [41] Parkes et al [29] Sharif et al [34] Wiadji et al [28]
<ul style="list-style-type: none"> Collect data on patient and staff feedback on the intervention for improvement purposes 	<ul style="list-style-type: none"> Jansson et al [41] Leppla et al [25]
<ul style="list-style-type: none"> Provide patients with tools to help assess, interpret, and act upon symptoms 	<ul style="list-style-type: none"> Leppla et al [25]

^aQoS: quality of service.

^bEMR: electronic medical record.

Category 4: Methods to Measure and Quantify the Impact of RPM on Clinical Staff

This category presents the methods used to identify the impact of RPM interventions on clinical staff tasks and workflows. In total, 2 main themes were established (Table 4) based on the kind of measures of the impact of RPM interventions on staff being collected and analyzed using different methods. The first theme was time and activity analysis, which includes methods for measuring clinical staff time expenditure and workload in relation to existing activities and RPM interventions. These

methods allow for a comparative analysis between the standard of care and the RPM intervention. Other possible quantifiable measures found in this category are the number of times certain resources are accessed or the time spent on certain tasks.

The second theme was staff satisfaction and experience, which focuses on how RPM interventions are perceived by the staff and how the new tools and ways of working affect their behaviors. This theme includes subjective measures, such as those gathered through interviews and surveys, and more objective measures, such as measures of adherence to protocols or alert frequency.

Table 4. Overview of methods to measure and quantify the impact of remote patient monitoring (RPM) interventions on clinical staff.

Theme and description	Studies
Time and activity analysis	
<ul style="list-style-type: none"> Activity timing <ul style="list-style-type: none"> Automatic recording of time spent on events and consultations Duration of use of RPM tools Cumulative time on activities Cumulative time on platform Frequency and quantity of alerts 	<ul style="list-style-type: none"> Downey et al [32] Makhni et al [31] Rothgangel et al [30] Sousa et al [37] Timmerman et al [27] Wiadji et al [28]
<ul style="list-style-type: none"> Activity mapping <ul style="list-style-type: none"> Current state mapping Implementation assessment Number of times telemonitoring tools were used Number of transmissions and events Selecting most busy times Nurses' tasks 	<ul style="list-style-type: none"> Augestad et al [38] Leppla et al [25] Rothgangel et al [30] Sousa et al [37] Timmerman et al [27]
<ul style="list-style-type: none"> Comparative analysis with baseline (time spent on activities and number of in-hospital visits and events) 	<ul style="list-style-type: none"> Harsha et al [21] Sousa et al [37]
<ul style="list-style-type: none"> Hospital logistics <ul style="list-style-type: none"> Number of in-hospital visits Length of in-hospital visits Type of complications Type of resources Accessibility to resources (quantity and quality) 	<ul style="list-style-type: none"> Augestad et al [38] Downey et al [32] Rothgangel et al [30]
<ul style="list-style-type: none"> Cost savings based on time and resources used 	<ul style="list-style-type: none"> Makhni et al [31]
Satisfaction and experience analysis	
<ul style="list-style-type: none"> Surveys and questionnaires <ul style="list-style-type: none"> Usability (eg, System Usability Score) Adherence to protocols Utility and efficiency of e-tools (frequency of incomplete data and effort and work needed for gathering extra data) 	<ul style="list-style-type: none"> Downey et al [32] Leppla et al [25] McMullen et al [33] Parkes et al [29] Rothgangel et al [30] Timmerman et al [27] Wiadji et al [28]
<ul style="list-style-type: none"> Interviews and focus groups 	<ul style="list-style-type: none"> Das et al [22] Davoody and Hägglund [23] Downey et al [32] Dunphy et al [36] Jansson et al [41] Ke et al [24] Leppla et al [25] McMullen et al [33] Parkes et al [29] Sharif et al [34]
<ul style="list-style-type: none"> Ethnographic research <ul style="list-style-type: none"> Observation Journey mapping 	<ul style="list-style-type: none"> Augestad et al [38] Das et al [22] Leppla et al [25] McMullen et al [33]
<ul style="list-style-type: none"> Co-design and cocreation sessions and workshops <ul style="list-style-type: none"> Critical incident technique—think-aloud approach—mock-ups 	<ul style="list-style-type: none"> Sanger et al [26] McMullen et al [33] Rothgangel et al [30]
<ul style="list-style-type: none"> Impact of alerts on performance and well-being 	<ul style="list-style-type: none"> Downey et al [32]

Discussion

Principal Findings

RPM is presented as a useful tool to help patients feel safer and more empowered in their self-care during the perioperative period. In addition, health care institutions benefit from it by increasing the efficiency in the use of their resources, both physical (such as beds and monitors) and human (clinical staff). In deciding on the adoption of RPM interventions, considering the impact on and perceptions of clinical staff is crucial as the success of these interventions is based on their cooperation and comprehension. As users and providers of remote perioperative care, clinical staff need to be comfortable and willing to adopt RPM interventions, which should not hinder their other tasks.

Overall, the main RPM-related problems found for clinical staff were related to undesirable changes in their workflow and lack of planning. In several included papers (11/21, 52%), the introduction of RPM led to a higher workload because of unforeseen tasks that emerged when the RPM intervention was implemented in the complex health care environment and not necessarily when the intervention was tested in controlled settings. In particular, tasks such as (remotely monitored) patient data analysis, remote alert response, and remote care reporting and billing were mentioned as sources of increased staff workload and disruptions in the usual care workflow. In addition, the time necessary for activities was often underestimated because of the lack of experience and knowledge of the clinical staff to perform some of the new tasks that RPM interventions created. Furthermore, problems were reported in relation to unanticipated users as sometimes it was unclear who was in charge of these new tasks, the assigned actor was not the adequate one, or they depended on the assistance of a third party. Problems regarding the difficulty in use and functioning of RPM tools were also described. This was mainly due to lack of knowledge or training, technical malfunctions, or legal issues where the new services conflicted with the current systems. Although it is true that these problems might be temporary and limited to the initial introduction of RPM interventions, it is still important to assess and address them as they do have an impact on the workflow and might cause the intervention implementation to fail before familiarization and adaptation are even possible. Furthermore, it is important to consider initial workflow problems as adaptation strategies and coping mechanisms adopted by staff to overcome these problems might in themselves generate structural issues. For example, when new tasks are introduced by RPM interventions without a clear indication of who is responsible for them, the available actors will feel compelled to take over, adding to their daily workload.

Most of the reported benefits for clinical staff related to the improvement in monitoring and data analysis, resulting in better resource management and clinical outcomes. Even though most staff members agree on the advantages these interventions bring in terms of better follow-up of patients and resource allocation, they are still concerned about the extra workload they face.

Regarding best practices and risk-reduction strategies, most of the included studies (18/21, 86%) mentioned the need to

strengthen the implementation process of RPM interventions through better planning and improved stakeholder involvement. This way, clinical staff can provide a better overview of their pre-existing work routines and needs so that the new interventions can be better integrated and adapted to their usual workflow rather than the other way around. Other strategies involved establishing protocols to guide the interventions' use and operations and providing the necessary training to avoid uncertainty and prepare staff. Finally, several included studies (10/21, 48%) stressed the importance of interoperability and ensuring compatibility between the new RPM interventions and the existing tools and processes used by the staff to prevent double work or the emergence of conflicts in the recorded patient data.

Moreover, it is recommended that RPM-related interfaces be user-friendly and tested in the context to reduce time spent on training and possible technical problems. Enhancing staff's understanding of and familiarity with the tools can increase their willingness toward their adoption as technology will be perceived as an enhancer and not as an obstacle.

The included studies reported recommendations for best practices and risk-reduction strategies for most of the staff-related problems and challenges mentioned in connection with RPM interventions. It is important to note that these solutions address problems that represent major barriers to RPM implementation in the present. Therefore, adopting them more consistently in RPM research and practice represents a way to maximize the capability of RPM to deliver real-world results in health care services in the future.

Figure 2 shows the connections between themes and categories. Here, we can see how some of the identified themes were not present in all the categories. Notably, there are problems that lack specific recommendations in the literature, such as those related to health care resources. Reimbursement schemes prioritizing in-hospital care constitute a largely unaddressed challenge complicated by the complexity of the context and the different types of stakeholders involved. This affects the commitment and motivation of clinical staff toward RPM interventions as it is not always clear how the extra or new work will be reimbursed. In addition, there are currently few answers on how to increase funding for RPM interventions (Table 1). This is a big challenge, as RPM interventions may not clearly present benefits justifying their relatively high expenses, especially in the short term.

There is still room for improvement in ways to manage incoming alerts so that they do not create interruptions and annoyance among staff while ensuring timely responses. Another open challenge is related to providing a collaborative environment between the different staff members involved in patient care and defining clear roles so as to divide RPM-related tasks effectively and avoid confusion. In addition, there are opportunities to improve the devices and systems that collect, analyze, and communicate patient data. This includes the possibility of using data for more informed or automatized decisions that consider multiple data sources, thus avoiding biases, false positives, and incorrect inputs.

Most of the methods used to assess the impact of RPM interventions on staff-related workflows were qualitative and subjective, including interviews, questionnaires, and observations. Few reported studies (7/21, 33%) included the collection of quantitative measures such as tracking the time invested in using the interventions. This is characterized as an opportunity for improvement in RPM-related research as quantitative impact measures would help assess resource use and, therefore, better evaluate the overall interventions. Furthermore, quantitative measures could unlock the possibility of meaningful comparisons across different interventions and contexts. Some of these more quantitative or objective measures could be anxiety levels using existing scales, as proposed by Jukic et al [42].

However, there is still not enough research on methods to track RPM-related workload quantitatively. Examples of RPM interventions in fields other than perioperative care can be useful in this regard. For example, in tele-intensive care units and the remote monitoring of cardiovascular implantable electronic devices, diverse methods have been deployed to measure staff workload [43-46] by, for example, time-motion studies [47,48]. In these interventions, systems automatically record use time while an observer also tracks the nurses and annotates the duration of RPM-related activities. This has reportedly helped researchers identify the most time-consuming aspects of RPM-related workflow and find bottlenecks and weaknesses to improve designs and implementation plans. These tele-intensive care unit and cardiovascular implantable electronic device remote monitoring research methods could be profitably translated to perioperative care. In general, research on RPM interventions [31,35] helps in understanding possible outcomes and identifying barriers, facilitators, and recommendations [30], which can guide the design and implementation stages of these interventions.

Further research should be dedicated to the quantification of resource use in RPM interventions—to standardize reimbursement policies—and the evaluation of the implementation of these strategies in different settings. Moreover, the time horizon of these studies should be extended to cover longer periods, as many relevant effects of RPM interventions cannot be observed in the short term - partly because of factors such as the staff learning curve.

Limitations

This scoping review has several limitations. The first is the diversity of RPM interventions examined as they might have different objectives, leading to variable results and problems.

In addition, the results will be influenced by the initial state and environment in which the RPM intervention was introduced. As mentioned by Herdman [49], intervention benefits depend on the baseline, whereby an initial higher performance may lead to a comparatively smaller advantage. Moreover, these interventions were executed under different circumstances and environments, which might change the dynamics among the clinical staff. Additional limitations are derived from the differences in the methodology used in the included studies as the target variables and outcomes might not be comparable. Finally, most of the included studies (13/21, 62%) only considered short- and midterm impacts, whereas RPM interventions can have long-term effects that are decisive to assess their overall performance.

This review was also susceptible to risk of bias because of missing results. This risk is increased by our exclusive focus on articles in English, our use of 2 databases (PubMed and Embase), and our focus on a limited time frame (January 2015 to March 2021). Nonreporting bias risk is also likely to apply to this review as we noticed that only a small fraction of papers in the RPM domain reported any insight at all on the impact of the introduction of new interventions on staff workflow. Overall, in light of the aforementioned limitations and risks of bias, we recommend interpreting and using our contribution as an initial description of the types of workflow-related implications of RPM described in the current literature and not as an exhaustive overview.

Conclusions

Every day there are more studies that show the impact of RPM interventions given their increasing use in clinical practice and in perioperative care in particular. Most of these studies focus on the patient's perspective and on clinical outcomes. In our scoping review, we presented an overview of the recent knowledge regarding clinical staff's perspective, which reveals the possible problems and benefits that remote monitoring can bring. Further research regarding policy making and protocol standardization should be conducted to establish a more trustworthy analysis of RPM interventions.

Studies concerning the impact of RPM strategies on clinical staff workflows and dynamics should be clear about the study objective, the design, and the methods used to test the intervention. This will help future readers in assessing the overall performance of RPM interventions. Moreover, this can enable better comparative research and promote the establishment of valuable benchmarking and auditing systems.

Conflicts of Interest

None declared.

Multimedia Appendix 1

Search strategy.

[[DOCX File, 15 KB - humanfactors_v9i2e37204_app1.docx](#)]

References

1. de Farias FA, Dagostini CC, de Assunção Bicca Y, Falavigna VF, Falavigna A. Remote patient monitoring: a systematic review. *Telemed J E Health* 2020 May;26(5):576-583. [doi: [10.1089/tmj.2019.0066](https://doi.org/10.1089/tmj.2019.0066)] [Medline: [31314689](https://pubmed.ncbi.nlm.nih.gov/31314689/)]
2. Seto E. Cost comparison between telemonitoring and usual care of heart failure: a systematic review. *Telemed J E Health* 2008 Sep;14(7):679-686. [doi: [10.1089/tmj.2007.0114](https://doi.org/10.1089/tmj.2007.0114)] [Medline: [18817497](https://pubmed.ncbi.nlm.nih.gov/18817497/)]
3. Koopman RJ, Wakefield BJ, Johanning JL, Keplinger LE, Kruse RL, Bomar M, et al. Implementing home blood glucose and blood pressure telemonitoring in primary care practices for patients with diabetes: lessons learned. *Telemed J E Health* 2014 Mar;20(3):253-260 [FREE Full text] [doi: [10.1089/tmj.2013.0188](https://doi.org/10.1089/tmj.2013.0188)] [Medline: [24350806](https://pubmed.ncbi.nlm.nih.gov/24350806/)]
4. van Os-Medendorp H, Koffijberg H, Eland-de Kok PC, van der Zalm A, de Bruin-Weller MS, Pasmans SG, et al. E-health in caring for patients with atopic dermatitis: a randomized controlled cost-effectiveness study of Internet-guided monitoring and online self-management training. *Br J Dermatol* 2012 May;166(5):1060-1068. [doi: [10.1111/j.1365-2133.2012.10829.x](https://doi.org/10.1111/j.1365-2133.2012.10829.x)] [Medline: [22268960](https://pubmed.ncbi.nlm.nih.gov/22268960/)]
5. Ricci RP, Morichelli L, Santini M. Remote control of implanted devices through Home Monitoring technology improves detection and clinical management of atrial fibrillation. *Europace* 2009 Jan;11(1):54-61. [doi: [10.1093/europace/eun303](https://doi.org/10.1093/europace/eun303)] [Medline: [19011260](https://pubmed.ncbi.nlm.nih.gov/19011260/)]
6. Anthony Jnr B. Use of telemedicine and virtual care for remote treatment in response to COVID-19 pandemic. *J Med Syst* 2020 Jun 15;44(7):132 [FREE Full text] [doi: [10.1007/s10916-020-01596-5](https://doi.org/10.1007/s10916-020-01596-5)] [Medline: [32542571](https://pubmed.ncbi.nlm.nih.gov/32542571/)]
7. Atilgan K, Onuk BE, Köksal Coşkun P, Yeşi LG, Aslan C, Çolak A, et al. Remote patient monitoring after cardiac surgery: the utility of a novel telemedicine system. *J Card Surg* 2021 Nov;36(11):4226-4234. [doi: [10.1111/jocs.15962](https://doi.org/10.1111/jocs.15962)] [Medline: [34478205](https://pubmed.ncbi.nlm.nih.gov/34478205/)]
8. Belarmino A, Walsh R, Alshak M, Patel N, Wu R, Hu J. Feasibility of a mobile health application to monitor recovery and patient-reported outcomes after robot-assisted radical prostatectomy. *Eur Urol Oncol* 2019 Jul;2(4):425-428. [doi: [10.1016/j.euo.2018.08.016](https://doi.org/10.1016/j.euo.2018.08.016)] [Medline: [31277778](https://pubmed.ncbi.nlm.nih.gov/31277778/)]
9. Dirnberger J, Waisbren S. Efficacy of telehealth visits for postoperative care at the Minneapolis VA. *Am J Surg* 2020 Sep;220(3):721-724. [doi: [10.1016/j.amjsurg.2020.01.015](https://doi.org/10.1016/j.amjsurg.2020.01.015)] [Medline: [31980135](https://pubmed.ncbi.nlm.nih.gov/31980135/)]
10. Symer MM, Abelson JS, Milsom J, McClure B, Yeo HL. A mobile health application to track patients after gastrointestinal surgery: results from a pilot study. *J Gastrointest Surg* 2017 Sep;21(9):1500-1505. [doi: [10.1007/s11605-017-3482-2](https://doi.org/10.1007/s11605-017-3482-2)] [Medline: [28685388](https://pubmed.ncbi.nlm.nih.gov/28685388/)]
11. van der Meij E, Huirne JA, Ten Cate AD, Stockmann HB, Scholten PC, Davids PH, et al. A perioperative eHealth program to enhance postoperative recovery after abdominal surgery: process evaluation of a randomized controlled trial. *J Med Internet Res* 2018 Jan 02;20(1):e1 [FREE Full text] [doi: [10.2196/jmir.8338](https://doi.org/10.2196/jmir.8338)] [Medline: [29295808](https://pubmed.ncbi.nlm.nih.gov/29295808/)]
12. Baniyasi T, Ghazisaedi M, Hassaniazad M, Niakan Kalhori SR, Shahi M, Esmaeili M. Surgical patients follow-up by smartphone-based applications: a systematic literature review. *Stud Health Technol Inform* 2020 Jun 23;271:85-92. [doi: [10.3233/SHTI200079](https://doi.org/10.3233/SHTI200079)] [Medline: [32578546](https://pubmed.ncbi.nlm.nih.gov/32578546/)]
13. Mehta SJ, Hume E, Troxel AB, Reitz C, Norton L, Lacko H, et al. Effect of remote monitoring on discharge to home, return to activity, and rehospitalization after hip and knee arthroplasty: a randomized clinical trial. *JAMA Netw Open* 2020 Dec 01;3(12):e2028328 [FREE Full text] [doi: [10.1001/jamanetworkopen.2020.28328](https://doi.org/10.1001/jamanetworkopen.2020.28328)] [Medline: [33346847](https://pubmed.ncbi.nlm.nih.gov/33346847/)]
14. Shah M, Douglas J, Carey R, Daftari M, Smink T, Paisley A, et al. Reducing ER visits and readmissions after head and neck surgery through a phone-based quality improvement program. *Ann Otol Rhinol Laryngol* 2021 Jan;130(1):24-31. [doi: [10.1177/0003489420937044](https://doi.org/10.1177/0003489420937044)] [Medline: [32567391](https://pubmed.ncbi.nlm.nih.gov/32567391/)]
15. Viers BR, Lightner DJ, Rivera ME, Tollefson MK, Boorjian SA, Karnes RJ, et al. Efficiency, satisfaction, and costs for remote video visits following radical prostatectomy: a randomized controlled trial. *Eur Urol* 2015 Oct;68(4):729-735. [doi: [10.1016/j.eururo.2015.04.002](https://doi.org/10.1016/j.eururo.2015.04.002)] [Medline: [25900782](https://pubmed.ncbi.nlm.nih.gov/25900782/)]
16. Forbes RC, Rybacki DB, Johnson TB, Hannah-Gillis A, Shaffer D, Hale DA. A cost comparison for telehealth utilization in the kidney transplant waitlist evaluation process. *Transplantation* 2018 Feb;102(2):279-283. [doi: [10.1097/TP.0000000000001903](https://doi.org/10.1097/TP.0000000000001903)] [Medline: [28767534](https://pubmed.ncbi.nlm.nih.gov/28767534/)]
17. Carayon P, Cartmill R, Hoonakker P, Hundt AS, Karsh BT, Krueger D, et al. Human factors analysis of Work™ow in health information technology implementation. In: Carayon P, editor. *Handbook of Human Factors and Ergonomics in Health Care and Patient Safety*. 2nd ed. Boca Raton, FL, USA: CRC Press; 2016:536-551.
18. Hignett S, Jones EL, Miller D, Wolf L, Modi C, Shahzad MW, et al. Human factors and ergonomics and quality improvement science: integrating approaches for safety in healthcare. *BMJ Qual Saf* 2015 Apr;24(4):250-254 [FREE Full text] [doi: [10.1136/bmjqs-2014-003623](https://doi.org/10.1136/bmjqs-2014-003623)] [Medline: [25715799](https://pubmed.ncbi.nlm.nih.gov/25715799/)]
19. Tricco AC, Lillie E, Zarin W, O'Brien KK, Colquhoun H, Levac D, et al. PRISMA extension for scoping reviews (PRISMA-ScR): checklist and explanation. *Ann Intern Med* 2018 Oct 02;169(7):467-473 [FREE Full text] [doi: [10.7326/M18-0850](https://doi.org/10.7326/M18-0850)] [Medline: [30178033](https://pubmed.ncbi.nlm.nih.gov/30178033/)]
20. Page MJ, Higgins JP, Sterne JA. Chapter 13: Assessing risk of bias due to missing results in a synthesis. In: Higgins JP, Thomas JA, editors. *Cochrane Handbook for Systematic Reviews of Interventions*. Version 6.3. London, UK: Cochrane; 2022.

21. Harsha P, Paul JE, Chong MA, Buckley N, Tidy A, Clarke A, et al. Challenges with continuous pulse oximetry monitoring and wireless clinician notification systems after surgery: reactive analysis of a randomized controlled trial. *JMIR Med Inform* 2019 Oct 28;7(4):e14603 [FREE Full text] [doi: [10.2196/14603](https://doi.org/10.2196/14603)] [Medline: [31661079](https://pubmed.ncbi.nlm.nih.gov/31661079/)]
22. Das A, Faxvaag A, Svanæs D. The impact of an eHealth portal on health care professionals' interaction with patients: qualitative study. *J Med Internet Res* 2015 Nov 24;17(11):e267 [FREE Full text] [doi: [10.2196/jmir.4950](https://doi.org/10.2196/jmir.4950)] [Medline: [26601678](https://pubmed.ncbi.nlm.nih.gov/26601678/)]
23. Davoody N, Hägglund M. Care Professionals' perceived usefulness of eHealth for post-discharge stroke patients. In: Hoerbst A, Hackl WO, de Keizer N, Prokosch HU, Hercigonja-Szekeres M, de Lusignan S, editors. *Exploring Complexity in Health: An Interdisciplinary Systems Approach*. Amsterdam, The Netherlands: IOS Press; 2016:589-593.
24. Ke JX, George RB, Wozney L, Chorney JL. Patient-centred perioperative mobile application in cesarean delivery: needs assessment and development. *Obstet Anesth Dig* 2020 Dec;40(4):185-186 [FREE Full text] [doi: [10.1097/01.aoa.0000719500.35419.e1](https://doi.org/10.1097/01.aoa.0000719500.35419.e1)]
25. Leppla L, Mielke J, Kunze M, Mauthner O, Teynor A, Valenta S, SMILe study team. Clinicians and patients perspectives on follow-up care and eHealth support after allogeneic hematopoietic stem cell transplantation: a mixed-methods contextual analysis as part of the SMILe study. *Eur J Oncol Nurs* 2020 Apr;45:101723. [doi: [10.1016/j.ejon.2020.101723](https://doi.org/10.1016/j.ejon.2020.101723)] [Medline: [32062362](https://pubmed.ncbi.nlm.nih.gov/32062362/)]
26. Sanger PC, Hartzler A, Lordon RJ, Armstrong CA, Lober WB, Evans HL, et al. A patient-centered system in a provider-centered world: challenges of incorporating post-discharge wound data into practice. *J Am Med Inform Assoc* 2016 May;23(3):514-525 [FREE Full text] [doi: [10.1093/jamia/ocv183](https://doi.org/10.1093/jamia/ocv183)] [Medline: [26977103](https://pubmed.ncbi.nlm.nih.gov/26977103/)]
27. Timmerman JG, Dekker-van Weering MG, Stuiver MM, Groen WG, Wouters MW, Tönis TM, et al. Ambulant monitoring and Web-accessible home-based exercise program during outpatient follow-up for resected lung cancer survivors: actual use and feasibility in clinical practice. *J Cancer Surviv* 2017 Dec;11(6):720-731 [FREE Full text] [doi: [10.1007/s11764-017-0611-6](https://doi.org/10.1007/s11764-017-0611-6)] [Medline: [28396993](https://pubmed.ncbi.nlm.nih.gov/28396993/)]
28. Wiadji E, Mackenzie L, Reeder P, Gani JS, Carroll R, Smith S, et al. Utilization of telehealth by surgeons during the COVID 19 pandemic in Australia: lessons learnt. *ANZ J Surg* 2021 Apr;91(4):507-514 [FREE Full text] [doi: [10.1111/ans.16693](https://doi.org/10.1111/ans.16693)] [Medline: [33634949](https://pubmed.ncbi.nlm.nih.gov/33634949/)]
29. Parkes RJ, Palmer J, Wingham J, Williams DH. Is virtual clinic follow-up of hip and knee joint replacement acceptable to patients and clinicians? A sequential mixed methods evaluation. *BMJ Open Qual* 2019 Mar 1;8(1):e000502. [doi: [10.1136/bmjopen-2018-000502](https://doi.org/10.1136/bmjopen-2018-000502)] [Medline: [31259271](https://pubmed.ncbi.nlm.nih.gov/31259271/)]
30. Rothgangel A, Kanera I, van den Heuvel R, Wientgen M, Jamin G, Lenssen T, et al. Physiotherapists' clinical use and acceptance of a telemonitoring platform during anterior cruciate ligament rehabilitation: a prospective clinical study. *Disabil Rehabil Assist Technol* 2022 Feb;17(2):184-191. [doi: [10.1080/17483107.2020.1774810](https://doi.org/10.1080/17483107.2020.1774810)] [Medline: [32521183](https://pubmed.ncbi.nlm.nih.gov/32521183/)]
31. Makhni MC, Riew GJ, Sumathipala MG. Telemedicine in orthopaedic surgery: challenges and opportunities. *J Bone Joint Surg Am* 2020 Jul 01;102(13):1109-1115. [doi: [10.2106/JBJS.20.00452](https://doi.org/10.2106/JBJS.20.00452)] [Medline: [32618908](https://pubmed.ncbi.nlm.nih.gov/32618908/)]
32. Downey CL, Croft J, Buckley H, Randell R, Brown JM, Jayne DG. Trial of Remote Continuous versus Intermittent NEWS monitoring after major surgery (TRaCINg): protocol for a feasibility randomised controlled trial. *Pilot Feasibility Stud* 2018 Jun 11;4:112 [FREE Full text] [doi: [10.1186/s40814-018-0299-z](https://doi.org/10.1186/s40814-018-0299-z)] [Medline: [29992041](https://pubmed.ncbi.nlm.nih.gov/29992041/)]
33. McMullen C, Nielsen M, Firemark A, Price PM, Nakatani D, Tuthill J, et al. Designing for impact: identifying stakeholder-driven interventions to support recovery after major cancer surgery. *Support Care Cancer* 2018 Dec;26(12):4067-4076 [FREE Full text] [doi: [10.1007/s00520-018-4276-0](https://doi.org/10.1007/s00520-018-4276-0)] [Medline: [29876832](https://pubmed.ncbi.nlm.nih.gov/29876832/)]
34. Sharif F, Rahman A, Tonner E, Ahmed H, Haq I, Abbas R, et al. Can technology optimise the pre-operative pathway for elective hip and knee replacement surgery: a qualitative study. *Perioper Med (Lond)* 2020 Nov 16;9(1):33 [FREE Full text] [doi: [10.1186/s13741-020-00166-0](https://doi.org/10.1186/s13741-020-00166-0)] [Medline: [33292556](https://pubmed.ncbi.nlm.nih.gov/33292556/)]
35. Brophy PD. Overview on the challenges and benefits of using telehealth tools in a pediatric population. *Adv Chronic Kidney Dis* 2017 Jan;24(1):17-21. [doi: [10.1053/j.ackd.2016.12.003](https://doi.org/10.1053/j.ackd.2016.12.003)] [Medline: [28224938](https://pubmed.ncbi.nlm.nih.gov/28224938/)]
36. Dunphy E, Hamilton FL, Spasić I, Button K. Acceptability of a digital health intervention alongside physiotherapy to support patients following anterior cruciate ligament reconstruction. *BMC Musculoskelet Disord* 2017 Nov 21;18(1):471 [FREE Full text] [doi: [10.1186/s12891-017-1846-0](https://doi.org/10.1186/s12891-017-1846-0)] [Medline: [29162071](https://pubmed.ncbi.nlm.nih.gov/29162071/)]
37. Sousa P, OLiveira M, Sousa J. Impact of computerized information systems in the global nurses' workload: nurses' perceptions and real-time. In: *Proceedings of the 3rd IPLeia's International Health Congress*. 2016 Presented at: IPLeia-IHC '16; May 6-7, 2016; Leiria, Portugal p. 27.
38. Augestad KM, Sneve AM, Lindsetmo RO. Telemedicine in postoperative follow-up of STOMa Patients: a randomized clinical trial (the STOMPA trial). *Br J Surg* 2020 Apr;107(5):509-518. [doi: [10.1002/bjs.11491](https://doi.org/10.1002/bjs.11491)] [Medline: [32100297](https://pubmed.ncbi.nlm.nih.gov/32100297/)]
39. Richards HS, Blazeby JM, Portal A, Harding R, Reed T, Lander T, et al. A real-time electronic symptom monitoring system for patients after discharge following surgery: a pilot study in cancer-related surgery. *BMC Cancer* 2020 Jun 10;20(1):543 [FREE Full text] [doi: [10.1186/s12885-020-07027-5](https://doi.org/10.1186/s12885-020-07027-5)] [Medline: [32522163](https://pubmed.ncbi.nlm.nih.gov/32522163/)]
40. Semple JL, Evans HL, Lober WB, Lavalley DC. Implementing mobile health interventions to capture post-operative patient-generated health data. *Surg Infect (Larchmt)* 2019 Oct;20(7):566-570. [doi: [10.1089/sur.2019.151](https://doi.org/10.1089/sur.2019.151)] [Medline: [31429637](https://pubmed.ncbi.nlm.nih.gov/31429637/)]

41. Jansson MM, Harjumaa M, Puhto AP, Pikkarainen M. Healthcare professionals' proposed eHealth needs in elective primary fast-track hip and knee arthroplasty journey: a qualitative interview study. *J Clin Nurs* 2019 Dec;28(23-24):4434-4446. [doi: [10.1111/jocn.15028](https://doi.org/10.1111/jocn.15028)] [Medline: [31408555](https://pubmed.ncbi.nlm.nih.gov/31408555/)]
42. Jukic T, Ihan A, Petek Šter M, Strojnik V, Stubljarić D, Starc A. Adherence of female health care workers to the use of a Web-based tool for improving and modifying lifestyle: prospective target group pilot study. *J Med Internet Res* 2020 Aug 14;22(8):e19500 [FREE Full text] [doi: [10.2196/19500](https://doi.org/10.2196/19500)] [Medline: [32687475](https://pubmed.ncbi.nlm.nih.gov/32687475/)]
43. Cady R, Finkelstein S, Lindgren B, Robiner W, Lindquist R, VanWormer A, et al. Exploring the translational impact of a home telemonitoring intervention using time-motion study. *Telemed J E Health* 2010 Jun;16(5):576-584 [FREE Full text] [doi: [10.1089/tmj.2009.0148](https://doi.org/10.1089/tmj.2009.0148)] [Medline: [20575725](https://pubmed.ncbi.nlm.nih.gov/20575725/)]
44. Papavasileiou LP, Forleo GB, Panattoni G, Schirripa V, Minni V, Magliano G, et al. Work burden with remote monitoring of implantable cardioverter defibrillator: is it time for reimbursement policies? *J Cardiovasc Med (Hagerstown)* 2013 Feb;14(2):114-119. [doi: [10.2459/JCM.0b013e328354e3e1](https://doi.org/10.2459/JCM.0b013e328354e3e1)] [Medline: [22644407](https://pubmed.ncbi.nlm.nih.gov/22644407/)]
45. Ricci RP, Morichelli L, D'Onofrio A, Calò L, Vaccari D, Zanotto G, et al. Manpower and outpatient clinic workload for remote monitoring of patients with cardiac implantable electronic devices: data from the HomeGuide Registry. *J Cardiovasc Electrophysiol* 2014 Nov;25(11):1216-1223. [doi: [10.1111/jce.12482](https://doi.org/10.1111/jce.12482)] [Medline: [24964380](https://pubmed.ncbi.nlm.nih.gov/24964380/)]
46. Ricci RP, Morichelli L, D'Onofrio A, Calò L, Vaccari D, Zanotto G, et al. Effectiveness of remote monitoring of CIEDs in detection and treatment of clinical and device-related cardiovascular events in daily practice: the HomeGuide Registry. *Europace* 2013 Jul;15(7):970-977 [FREE Full text] [doi: [10.1093/europace/eus440](https://doi.org/10.1093/europace/eus440)] [Medline: [23362021](https://pubmed.ncbi.nlm.nih.gov/23362021/)]
47. Tang Z, Mazabob J, Weavind L, Thomas E, Johnson TR. A time-motion study of registered nurses' workflow in intensive care unit remote monitoring. *AMIA Annu Symp Proc* 2006;2006:759-763 [FREE Full text] [Medline: [17238443](https://pubmed.ncbi.nlm.nih.gov/17238443/)]
48. Tang Z, Weavind L, Mazabob J, Thomas EJ, Chu-Weininger MY, Johnson TR. Workflow in intensive care unit remote monitoring: a time-and-motion study. *Crit Care Med* 2007 Sep;35(9):2057-2063. [doi: [10.1097/01.ccm.0000281516.84767.96](https://doi.org/10.1097/01.ccm.0000281516.84767.96)] [Medline: [17855819](https://pubmed.ncbi.nlm.nih.gov/17855819/)]
49. Herdman TH, Burgess LP, Ebright PR, Paulson SS, Powell-Cope G, Hancock H, et al. Impact of continuous vigilance monitoring on nursing workflow. *J Nurs Adm* 2009 Mar;39(3):123-129. [doi: [10.1097/NNA.0b013e31819894e8](https://doi.org/10.1097/NNA.0b013e31819894e8)] [Medline: [19590468](https://pubmed.ncbi.nlm.nih.gov/19590468/)]

Abbreviations

PRISMA-ScR: Preferred Reporting Items for Systematic Reviews and Meta-Analyses extension for Scoping Reviews

RPM: remote patient monitoring

Edited by A Kushniruk; submitted 10.02.22; peer-reviewed by H Karim, B Zhao; comments to author 21.03.22; revised version received 04.04.22; accepted 23.04.22; published 06.06.22.

Please cite as:

León MA, Pannunzio V, Kleinsmann M

The Impact of Perioperative Remote Patient Monitoring on Clinical Staff Workflows: Scoping Review

JMIR Hum Factors 2022;9(2):e37204

URL: <https://humanfactors.jmir.org/2022/2/e37204>

doi: [10.2196/37204](https://doi.org/10.2196/37204)

PMID: [35666574](https://pubmed.ncbi.nlm.nih.gov/35666574/)

©Maria Alejandra León, Valeria Pannunzio, Maaike Kleinsmann. Originally published in *JMIR Human Factors* (<https://humanfactors.jmir.org>), 06.06.2022. This is an open-access article distributed under the terms of the Creative Commons Attribution License (<https://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in *JMIR Human Factors*, is properly cited. The complete bibliographic information, a link to the original publication on <https://humanfactors.jmir.org>, as well as this copyright and license information must be included.

Review

User Perceptions of eHealth and mHealth Services Promoting Physical Activity and Healthy Diets: Systematic Review

Julia Bergevi¹, MSc; Susanne Andermo^{2,3}, PhD; Yohannes Woldamanuel¹, MPH; Unn-Britt Johansson^{1,4}, PhD; Maria Hagströmer^{1,2,5}, PhD; Jenny Rossen¹, PhD

¹Department of Health Promoting Science, Sophiahemmet University, Stockholm, Sweden

²Department of Neurobiology, Care Sciences, and Society, Karolinska Institutet, Stockholm, Sweden

³Department of Public Health Sciences, Karolinska Institutet, Stockholm, Sweden

⁴Department of Clinical Sciences and Education, Karolinska Institutet, Stockholm, Sweden

⁵Academic Primary Health Care Center, Stockholm, Sweden

Corresponding Author:

Jenny Rossen, PhD

Department of Health Promoting Science

Sophiahemmet University

Lindstedsvägen 8

Stockholm, 11486

Sweden

Phone: 46 701718634

Email: jenny.rossen@shh.se

Abstract

Background: Physical activity and a diet that follows general recommendations can help to prevent noncommunicable diseases. However, most adults do not meet current recommended guidelines, and support for behavior change needs to be strengthened. There is growing evidence that shows the benefits of eHealth and mobile health (mHealth) services in promoting healthy habits; however, their long-term effectiveness is uncertain because of nonadherence.

Objective: We aimed to explore users' perceptions of acceptability, engagement, and usability of eHealth and mHealth services that promote physical activity, healthy diets, or both in the primary or secondary prevention of noncommunicable diseases.

Methods: We conducted a systematic review with a narrative synthesis. We performed the literature search in PubMed, PsycINFO, and CINAHL electronic databases in February 2021 and July 2021. The search was limited to papers published in English between 2016 and 2021. Papers on qualitative and mixed method studies that encompassed eHealth and mHealth services for adults with a focus on physical activity, healthy diet, or both in the primary or secondary prevention of noncommunicable diseases were included. Three authors screened the studies independently, and 2 of the authors separately performed thematic analysis of qualitative data.

Results: With an initial finding of 6308 articles and the removal of 427 duplicates, 23 articles were deemed eligible for inclusion in the review. Based on users' preferences, an overarching theme—eHealth and mHealth services provide value but need to be tailored to individual needs—and 5 subthemes—interactive and integrated; varying and multifunctional; easy, pedagogic, and attractive; individualized and customizable; and reliable—emerged.

Conclusions: New evidence on the optimization of digital services that promote physical activity and healthy diets has been synthesized. The findings represent users' perceptions of acceptability, engagement, and usability of eHealth and mHealth services and show that services should be personalized, dynamic, easily manageable, and reliable. These findings can help improve adherence to digital health-promoting services.

(*JMIR Hum Factors* 2022;9(2):e34278) doi:[10.2196/34278](https://doi.org/10.2196/34278)

KEYWORDS

acceptability; behavior change; engagement; health technology; noncommunicable diseases; usability; user feedback; qualitative studies; physical activity; healthy diet

Introduction

Noncommunicable diseases, such as type 2 diabetes, cardiovascular diseases, and certain types cancer (colon, breast, prostate), are the leading causes of impaired quality of life and premature death worldwide, responsible for 71% of all deaths globally [1]. In Europe, where 60% of incidences are associated with unhealthy lifestyles (such as poor diet and physical inactivity) [2], close to 800,000 EU citizens die yearly because of noncommunicable diseases. The noncommunicable disease epidemic continues to grow and is expected to cause 75% of all global deaths by 2030 [3]. World Health Organization guidelines on physical activity and sedentary behavior suggest that adults should perform at least 150 to 300 minutes of moderate-intensity aerobic exercise or 75 to 150 minutes of vigorous-intensity aerobic exercise per week [4]. If adults were more physically active, 4 to 5 million global deaths yearly could be prevented [4]. Yet, only 1 in 4 adults meet the global recommendations for physical activity [5]. There is also growing evidence that a healthy diet plays an important role in preventing noncommunicable diseases [6]. Dietary recommendations may vary between nations but originate from global guidelines [6] that suggest that adults should eat all macronutrients in balance with the energy expenditure; consume a limited amount of saturated fats, trans fats, sugars, and salt; and consume more fruits, vegetables, and whole grains.

eHealth has been defined as “the use of emerging information and communications technology to improve or enable health and health care [7].” A subsegment of eHealth is mobile health (mHealth), which has been defined as “medical and public health practice supported by mobile devices, such as mobile phones, patient monitoring devices, personal digital assistants, and other wireless devices [8].” eHealth has the potential to support behavior change, and thus, improve health. For instance, several studies that have investigated the effects of digital lifestyle interventions reported short-term positive effects in disease-specific clinical outcomes [9-11], physical activity levels [12,13], and dietary patterns [12,14]. There is also evidence that physical activity interventions delivered using technology are 12% more effective in increasing physical activity levels than those that are not delivered using technology [15]. However, 75% of people who download smartphone health apps stop using the apps within a short time [16]. There is a need to identify factors which influence engagement with and adherence to health-promoting technology [10,11,17,18].

Previous reviews of qualitative studies have captured users’ perceptions and beliefs about mHealth apps [19] or analyzed different behavior change techniques and persuasive system

designs in concern of users’ motivation and maintenance in eHealth tools [20]. However, to our knowledge there is no summarized evidence on users’ perceptions of factors that may affect the acceptability, engagement, and usability in eHealth and mHealth services that focus exclusively on physical activity, diet, and lifestyle-related diseases in the primary and secondary prevention of noncommunicable diseases. Filling this gap is vital to capitalize on the promising prospects of health technology. Therefore, this systematic review explores users’ perceptions of acceptability, engagement, and usability of eHealth and mHealth services that promote physical activity, healthy diets, or both in the primary or secondary prevention of noncommunicable diseases.

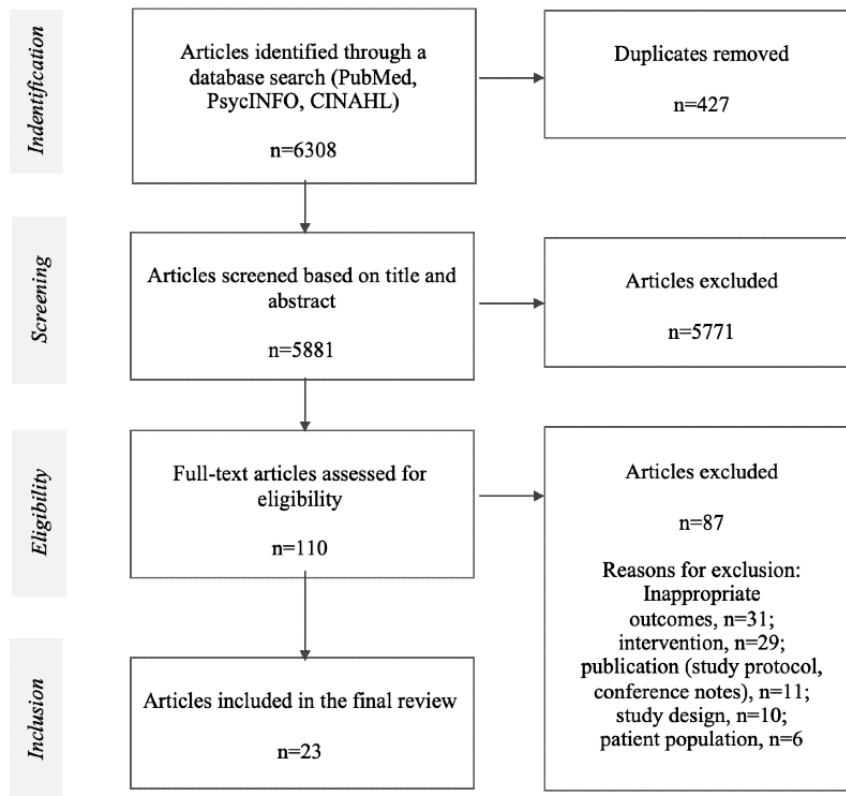
Methods

Overview

In this systematic review, qualitative studies were summarized using a narrative synthesis [21]. The process followed the PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-analyses) framework [22] (checklist [23] in [Multimedia Appendix 1](#)). Only studies with ethical approval were included to avoid encouraging eHealth interventions in which study participants may have harmed their physical or mental health. The review was registered on July 25, 2021 (PROSPERO International Prospective Register of Systematic Reviews; CRD42021261844).

Search Strategy

Assisted by 2 university librarians, we searched PubMed, PsycINFO, and CINAHL electronic databases in February 2021; we updated the search results in July 2021 ([Figure 1](#)). The search terms were (“acceptability” OR “engagement” OR “usability” OR) AND (“digital service” OR “eHealth” OR “mHealth”) AND (“behavior change” OR “physical activity” OR “diet”). A full overview of the search terms is listed in [Multimedia Appendix 2](#). The search was limited to papers published in English between 2016 and 2021, given that the rapidly progressing nature of health-promoting technology [24] likely lowered the relevance of older publications (ie, outdated technology). All identified studies were imported to review management software (Covidence systematic review software, Veritas Health Innovation) that automatically removed duplicates. Three authors (JB, YW, JR) independently screened titles and abstracts to determine whether papers would be included in the second screening phase. Any disagreements were discussed. In the second screening phase, the full texts were independently screened by 2 authors (JB, JR) to determine the final selection of papers.

Figure 1. Search strategy.

Selection Criteria

The selection criteria (Table 1) were based on the PEO (Population, Exposure, Outcome) framework [25]. We included papers describing qualitative or mixed methods studies that presented qualitative data on acceptability, engagement, and

the usability of digital behavior change services for the promotion of physical activity, healthy diets, or both, consistent with current World Health Organization guidelines [4,6]. We did not include papers that focused on interventions that targeted sedentary behaviors only.

Table 1. Selection criteria based on the PEO (Population, Exposure, Outcome) framework.

Criterion	Inclusion	Exclusion
Study type	Ethical approval, full text available, English language, all geographical locations, published between 2016 and 2021	No ethical approval, full text not available, non-English language, published before 2016
Study design	Original qualitative study, original or secondary analysis of a mixed methods study including a qualitative method	Systematic review, meta-analysis, study protocol, efficacy or effect study evaluating effect only
Population	Healthy adults (≥ 18 years) and adults with noncommunicable diseases, including overweight or obesity, type 2 diabetes, cardiovascular diseases and relevant cancer types (colon, breast, prostate)	Children and adolescents (≤ 18 years), pregnant women, clinical populations (eg, communicable diseases and severe diseases)
Exposure	Web-based platform or mobile app promoting lifestyle-related behavior changes on an individual level focusing on physical activity, healthy diets, or both in the primary or secondary prevention of noncommunicable diseases	Focus not explicitly on physical activity, healthy diets, or both (alcohol, tobacco, sleep, sedentary behavior, mental health, medical adherence), physical activity or diet not comparable to WHO ^a guidelines, behavior changes at group level (eg, group-oriented activities in workplace settings), content adapted to a specific clinical population (eg, cancer, cardiovascular diseases patients) and thus not suitable for general adults, platform or app including dietary recording, calorie counting, exergames, social networking, short message service, digital counseling, wearables only (ie, no multicomponent platforms or apps)
Outcome	Qualitative data on acceptability, engagement, and usability	Quantitative data on acceptability, engagement, and usability; qualitative data on a user's perceived effect or general experience in participating in a study; qualitative data on an individual's general preferences of eHealth technology (ie, no platform or app yet designed); health care professionals' perceptions; qualitative data only evaluating one feature of a platform or app

^aWHO: World Health Organization.

Data Extraction and Analysis

Data extraction was performed by JB and reviewed by JR in accordance with Guidance on the Conduct of Narrative Synthesis in Systematic Reviews [21]. The procedure included three steps: (1) tabulation to provide detailed information of all studies in a common table (Multimedia Appendix 3), including reference details, study design, population, exposure, outcome, and quality; (2) clustering to organize the findings into groups relevant to the research aim (participants' subjective opinion clustered as facilitators, barriers, and suggested improvements); and (3) translation to explore similarities and differences between the studies using thematic analysis (ie, identify the most relevant and important themes and concepts across the studies in an inductive manner, hence, without predefined themes to guide the analysis) [21]. The thematic synthesis included 3 steps, described by Thomas et al [26] as "coding of text line-by-line; the development of 'descriptive themes'; and the generation of 'analytical themes.'"

Quality Appraisal

Quality appraisal was performed in accordance with the process described in The Swedish National Agency for Medical and Social Evaluation method book [27]. Quality criteria (authors' affiliated departments, study design, study theory, recruitment, data collection, data analysis methodology, relevance to the study aim, coherence, sample size, and results) were assessed as high, medium, or low.

Results

Overview

A total of 6308 papers were identified; 427 duplicates were removed, 5881 titles and abstracts were screened for eligibility, and 110 full-text articles were read. The final sample comprised 23 papers (Figure 1).

Study Characteristics

Study characteristics are outlined in Multimedia Appendix 3. Papers were published between 2016 and 2021, with the majority published after 2018 ($n=18$). The studies were undertaken in the United Kingdom [28-33], Italy [34], Belgium [35,36], Finland [37,38], Portugal [39], Sweden [40], Germany [41], the United States [42-46], Australia [47-49], and Canada [50]. Of the included studies, 11 were qualitative studies [32,34,35,37,40,42-44,46,47,50], and 12 were mixed methods studies [28-31,33,36,38,39,41,45,48,49]. Most studies included participants of both sexes [28-41,43,44,46,47,49,50], except for 1 study [48] with men only, and 2 studies [42,45] with women only. One study did not report sex distribution [39]. In total, sex was reported for 417 women, 309 men, and 1 nonbinary individual. In the studies with both men and women, most often, women were overrepresented. In the 23 studies, 769 participants between 18 and 75 years (mean age range 34-62 years) were included. Ten studies included healthy adults [28,31,34-36,38-40,46,49,50], and 13 included adults with obesity or who were overweight [29,30,33,37,42,45,47,48], adults with type 2 diabetes [43], or adult cancer survivors [29,32,44].

More than half of the studies focused on apps [28,29,31,32,34,35,38,40,42,44,45] or websites [36,41] promoting physical activity, while the rest involved diet-promoting apps [30,33,49], or diet- and physical activity-promoting apps [39,43,46-48,50] or websites [37]. The most commonly applied behavior change techniques according to behavior change technique taxonomy [51] were (presented in descending order) feedback and monitoring (category 2) [28-31,33,35-42,44-48], goals and planning (category 1) [29-31,35,36,38,40,44,47,48,50], social support (category 3) [28,31,37,44,47,48], and rewards and threats (category 10) [29,31,33,38,44,48]. The services were tested for up to a 1-year period, with the majority (n=18) being tested for ≤ 4 months. One study [44] did not report the intervention period's duration. Of the 23 studies, 20 studies used semistructured interviews [29-41,44-50], 2 studies used think-aloud interviews [34,39], 5 studies used focus group discussions [30,33,42,43,45], and 3 studies used web-based questionnaires [39,41,49].

Study Quality

In all studies, a theoretical framework was used to support the purpose of the study; papers included information on authors with relevant professions, recruitment methods, data collection and analysis, and results: 10 studies were appraised as high quality [32,34,35,37,40,42,46,47,49,50], 11 studies were appraised as medium quality [29-31,36,38,39,41,43-45,48], and 2 studies were appraised as low quality [28,33] (Multimedia Appendix 3).

Findings

Overview

An overarching theme and 5 subthemes emerged from the thematic analysis reflecting users' perceived acceptability, engagement, and usability of eHealth and mHealth services promoting physical activity and healthy diets. The overarching theme showed that eHealth and mHealth services provide value but need to be tailored to create compelling services that offer long-term user value. The subthemes indicated that users prefer services to be (1) interactive and integrated; (2) varying and multifunctional; (3) easy, pedagogic, and attractive; (4) individualized and customizable; and (5) reliable.

Provide Value but Need to Be Tailored

Users recognize that eHealth and mHealth services can support behavior change but that more effective services are needed to meet individual needs, provide long-term user value and keep engagement over time.

Interactive and Integrated

Users stressed the importance of an interactive service, enhanced with a dynamic bidirectional communication path between the service and the user. Some users described dynamic communication as a desire to make services more human and less robotic, for instance, to make services operate as an automatic coach or to integrate services with a physical coach [35]. Several users expressed disappointment when the service was not sufficiently interactive or did not provide sufficient support [29,32,34-36,38,39,41,46,47]. More interactive guidance was expressly requested when including goal setting, action

planning, and coping planning [35,36,46,47]. The request for more interactive guidance was exemplified by 2 users who stated desire for and satisfaction with interactive guidance when discussing physical activity-promoting apps. One user said,

Something that gradually guide you toward your goals, step-by-step, perhaps also suggesting what kind of physical activity to do and providing advice. [34]

The other user remarked,

It provides suggestions about how much activity to do per week, how to increase it, etc. That what I liked a lot. [34]

Integration with health experts, external health devices, and support services to increase user engagement and usability was feedback commonly expressed by users. Some users wanted to connect with personal trainers, health coaches, and clinicians to receive information, recommendations, and feedback [32,34,37,48]. One study suggested inviting expert moderators to create more productive discussions when social networking [37]. Some studies reported that the integration of other device apps (eg, calendar, alarm, and external health apps) as suggestions for improvement [39,40,49]. Some users found it comfortable to track physical activity by phone [42,46], whereas others preferred the integration of an app with a wearable [29,31,34,35,47]. In addition, users appreciated automatic syncing with external apps or wearables that monitor multiple variables (eg, steps, distance, calories, heart rate) were integrated [35,40-42,47]. Some users proposed services that enable meal planning and food purchasing by integrating people's shopping lists with a web-based grocery service [49]. Other suggestions for integration included the ability to synchronize app content with family members and friends (eg, sharing goals and grocery lists), connect to sponsors that donate rewards when goals are achieved, obtain community resources and location-specific recommendations to facilitate physical activity, and arrange meetup-style events to gain support from and connect with peers online [29,42-44,47,49].

Varying and Multifunctional

Variety was another frequently cited theme of importance. One user stated,

If it always stays the same I think I will not use it for long and will consequently delete it [39]

The significance of variety applied both to the content and to the included behavior change techniques of the service. Users preferred variety or novelty over repetition for motivational, inspirational, and educational content [33,34,39-41,44,49]. One user said,

...it was the same exact wording in the message every single time, so it almost seemed like robotic. [44]

Apps with several behavior change techniques were appreciated. The behavior change techniques most appreciated by users were social networking [28-31,34-36,38-40,42-45,48], self-monitoring [28-31,33,34,37,38,40,43,47,48,50], push notifications [28-31, 34-37, 40, 44, 47], progress tracking [29,31,34-36,40,44,46], goal setting [34,40-42,44,47,50], and gamification (ie, gamified

challenges and rewards) [31,33,35,37,44,47,48]. Users also appreciated the ability to track several health parameters in the same app (eg, energy expenditure, heart rate, weight loss, physical activity level, diet, water intake) [34,40,44,47,50]. One user commented,

I would have liked an app that includes a wide variety of health measurements. Now there are apps for movement and apps for eating, but if you got them all in one app I would use it a lot more. If the app included other health components, I could have set goals that were more attractive to me. [40]

Yet, one study [38] found that services with too many options may be a hindrance for older people (≥ 63 years), which could potentially affect user acceptability and usability negatively. Another study [49] reported that it might be problematic to include too many features, as this could make the service difficult to navigate.

Easy, Pedagogic, and Attractive

Users recognized the value of a straightforward service with good flow and a menu that can be easily navigated [31,32,34-36,39,45-47,49,50]. At the same time, they disliked cognitively demanding or time-consuming services [30,34,43,48]. One user said,

I mean part of the reason why the step app worked so well was that you literally turn it on it does everything. There isn't really a lot I need to do to interact with it further. [46]

Another user noted,

Y'know...it's a nice, simple app. You don't need to be that literate. [32]

Users often reported manual data entry as an obstacle because it was time-consuming [34,35,37,43,48], especially in tracking physical activity or diet using a diary [37,41,43,48]. Users preferred easily performed exercises that did not require additional equipment [32,41]. In one study that used photos of meals for dietary self-monitoring, users found the method to be inappropriate in social settings [33]. Another study applied self-monitoring of food choices using "Happy-scores [39]." Users liked its easy and educational way to monitor and reflect on lifestyle habits. One user said,

We saw when we said we ate "bad" foods (fried food and such), and we lowered our score, it was...we thought "right, I shouldn't have eaten that" or "I should have eaten a healthier food." The fact that we have a score and we see the effect of that score in our behaviour ends up motivating us to have a better score. [39]

Visualization of goals and clinical parameters using easy-to-read graphs was either requested or appreciated as a way to track progress in a larger context [30,32,37,40,43,44,50]. In addition, many users wanted to be provided with a manual or initial tutorial to learn about the service or new tools. They also desired technical support [29,31,41,46,47,50]. Users valued the attractiveness of an app if the content and tone of the service are not discouraging or associated with illness and disease.

Overall, users preferred services to be encouraging, fun, and positive [29,37-40,44,45,49]. One user said,

Although [another sport app] it's just an app, but it says something like "now you've missed your training session," it makes me feel somehow bad. So probably you should pay attention to that, how the feedback is. [38]

What was perceived as an attractive layout varied widely among the users. While some preferred a clean design [31,35,36,46,50], others favored more color [35-37,48]. One study [42] stressed the importance of using a layout that was not too child-like (eg, excluding smiley faces) as it decrease the service's reliability. In 2 studies [47,50], changeable layout themes were offered, which the users appreciated. For external physical activity trackers, the users valued small, light, and waterproof devices [41,44].

Individualized and Customizable

Several studies [32,35-41,44-46,49,50] reported individualized content as a facilitator or suggestion for improvement when interacting with eHealth and mHealth apps. For example, users valued content tailored to personal motives and goals, current health status, fitness level, motivation level, season, weather conditions, and profile set-ups (such as sex, age, and personal interests) [32,34,36-38,40,41,50]. However, one study [38] reported concerns stereotyping based on interests or activities and emphasized the importance of modifiable individual set-ups. In one study [49], users suggested that recipes should be adapted to the family constellation (eg, modified portion sizes and meal suggestions appropriate to young children). For physical activity-promoting apps, some users noted the importance of offering relevant and challenging exercises [50]. Moreover, addressing users by their names was suggested (eg, when sending push notifications) [45]. Several studies [29,34,35,38,40,45] reported that users like to gain a sense of control of the service by customizing behavior change techniques to personal needs, preferences, and schedules. One user expressed discontent when the push notifications were not tailored to the person's schedule:

The amount of time is not much, but sometimes it is...because you get the notification at 8 o'clock, that didn't fit my working schedule. If I start with an early shift, I get up at 5 o'clock in the morning, at 6.30 o'clock I'm already at work...and then I actually have to think about my app during coffee break...And those things didn't always go so well... [35]

In addition, there were mixed opinions on certain elements. For instance, users did not agree with push notifications, social networking, and gamification: some appreciated or requested them [32,34,37,38,40], while others found them inappropriate or annoying [32,34,37,38,40,45]. Some users wanted to adjust push notifications to personal goals, frequency, and time [30,45]. Two studies also emphasized the importance of customizing the content to a user's self-identity (ie, sex, age, body size, and fitness level) when, for example, sharing activity tips using video clips and internet instructors [32,35]. One user said,

And of course, umm, on both of them [J&J and Gorilla Workout]...the videos, err, show the sort of slim, fit young, ultra-fit, young men doing it. You think, "Gosh, I...I haven't looked like that for about 40 years. [32]

Reliable

A reliable service, with proven personal safety and trustworthiness, was expressed as essential. Some users complained about sharing personal data and wanted confidentiality ensured before sharing private and sensitive data [32,43]. A service that originated from a trustful source (such as recognized authorities or health care professionals) and provided evidence-based content in line with public recommendations, was perceived as being more reliable [29,37,46,49,50]. User perception of the reliability of services decreased when excessive advertisements, when parts of the content were unavailable if not paid for, and regular system updates were part of the service [34,35,37,40,44,46]. When tracking the physical activity level by phone or wearable, users reported that it was important for the tracker to be convenient and technical accurate in distinguishing different activities (eg, walking, running, biking) [32,34,40]. Technical issues were generally perceived as impediments, with users expressing the need for apps to be technically stable, easily manageable, and effective. Finally, apps should not drain the battery, mobile data usage, or phone memory [32,44,49,50].

Discussion

Principal Results

In this systematic review, we explored adults' perceptions of the acceptability, engagement, and usability of eHealth and mHealth services that focus on physical activity, healthy diets, or both in the primary and secondary prevention of noncommunicable diseases. The results showed that users value eHealth and mHealth services, but considerations need to be taken account to maintain engagement. Users preferred services to be (1) interactive and integrated; (2) varying and multifunctional; (3) easy, pedagogic, and attractive; (4) individualized and customizable; and (5) reliable. By taking these findings into account, we believe that adherence to eHealth and mHealth services could be significantly improved.

Comparison With Previous Research

Users underlined the need for variation. This user view was supported by Dennison et al [52], who reasoned that new and updated content increases mHealth app users' motivation and engagement. Users also valued a service that is composed of several behavior change techniques. There was some disagreement about the effectiveness of behavior change techniques and about the number of behavior change techniques that should be employed. A meta-analysis [53] reported that intervention effectiveness increased when more behavior change techniques were included. In contrast, Kelders [54] underscored the importance of matching user and intervention characteristics rather than applying several behavior change techniques. Our findings indicate that users preferred individualized services. Users also valued a straightforward and easy-to-use service,

which is aligned with the *less-is-more* strategy for effective human–computer interactions [55]. A service offering multiple behavior change techniques tailored to users' preferences and activities can enhance user engagement. Broekhuizen et al [56] confirmed that tailoring to an individual's needs was beneficial in digital health behavior change interventions. On the other hand, tunneling the content and basing it on presumed stereotypical activities should be avoided as presumed assessments may mislead the tailoring process. Analytical and artificial intelligence–based methodologies that use input from the app or captured by external devices, could improve user individualization without increasing user burden [57]. Our findings also show that users valued an interactive service enhanced with bidirectional communication and indicative support, especially when the behavior change techniques were used for goals and planning (category 1 [51]). Evidence supports this finding, showing that it is crucial with indicative support to set realistic and achievable goals when minding motivation and engagement [58]. Some evidence stated the importance of applying well-established behavior change techniques when designing health-promoting technology, with suggestions to include self-monitoring and goal setting as support for physical activity and dietary behavior changes [59–61]. It is well evidenced that self-monitoring and goal setting appear to enhance the behavior change process and increase the intervention effect [59–61]. This review shows self-monitoring of several health parameters, goal setting, social networking, gamification, and push notifications were valued as behavior change techniques. Social networking, gamification, and push notifications were appreciated to gain support and enhance motivation, although some users found these behavior change techniques to be inappropriate or even annoying when not carefully adjusted to personal schedules, motives, and interests. There were conflicting opinions about social networking and gamification. For instance, some studies reported that social networking and gamification favored usability (ie, the efficacy and satisfaction of the service) [62,63], whereas others felt it was not essential for long-term behavior changes [18,64]. However, this difference in findings reflects users' individual preferences, which suggests that there is a need to offer a customizable and flexible device to provide a personalized and dynamic service that follows the varying attitudes, values, and schedules of users. Some users expressed privacy concerns when sharing personal data, which could be an issue when tailoring is used. Individualization and anonymity have been discussed as a problem in eHealth elsewhere [38,65]. Finally, our results suggest that time efficiency may be another crucial factor that is particularly challenging in monitoring dietary habits. This view was supported by Peng et al [65], who reported that ease of use and time efficiency was significant for long-term engagement to mHealth apps.

Strengths and Limitations

Our results are based on newly published studies, which is a strength given the rapid progression of eHealth and mHealth. Another strength is that we included studies with healthy adults and adults with a medical history and from a wide age range. Thus, the results can be generalized as users' perceptions may vary with age and the purpose of the service being used. We

also included studies with participants of both sexes and from several countries. However, a slightly larger number of women (417 women compared with 309 men) were included in this review, limiting the generalizability to male populations. Studies were from diverse high-income countries strengthening international generalization. Yet, few studies were from low- and middle-income countries which restricts generalization to low- and middle-income countries. Most of the studies included in this review had recruited participants interested in using health-promoting services. This recruitment bias could limit the generalizability of our results because such individuals are likely to have a level of motivation that is higher than that of the general population. Also, we included studies in which individuals used the digital service for free, which may affect the expectations and perceptions of the service compared with the real world, where consumers pay for services. Most studies had an intervention period of only 4 months, which is a limitation because users' perceptions are likely to change; however, one with a 12-month intervention period did not report

any deviating results. All studies were included regardless of assessed quality; however, none of the studies assessed with low quality added anything new or distinctive to the results. Also, this review only included qualitative studies; quantitatively measured aspects were not considered. Researcher bias is a potential limitation in analyzing qualitative data. However, this limitation is less of an issue, because in our review, 2 authors independently analyzed and discussed the findings.

Conclusion

Our findings from the synthesis of studies on the optimization of digital services to promote physical activity and healthy diets represent users' perceptions of acceptability, engagement, and usability and show that eHealth and mHealth services provide value but need to be tailored to make them personalized, dynamic, easily manageable, and reliable. These findings can be useful in improving the user value when receiving support by digital services for behavior change to promote healthy lifestyles and increase adherence to eHealth services.

Authors' Contributions

JB performed the literature analysis and wrote the draft of the manuscript. JR provided input on the literature analysis, helped combine the results from the literature analysis, assisted in writing, and reviewed and approved the paper. SA, YW, U-BJ, and MH provided input on the methods and results, reviewed and approved the draft, and provided critical feedback.

Conflicts of Interest

None declared.

Multimedia Appendix 1

PRISMA checklist.

[PDF File (Adobe PDF File), 4776 KB - [humanfactors_v9i2e34278_app1.pdf](#)]

Multimedia Appendix 2

Search terms.

[DOCX File , 55 KB - [humanfactors_v9i2e34278_app2.docx](#)]

Multimedia Appendix 3

Description of studies included in the review.

[DOCX File , 113 KB - [humanfactors_v9i2e34278_app3.docx](#)]

References

1. Noncommunicable diseases. World Health Organization. URL: <https://www.who.int/news-room/fact-sheets/detail/noncommunicable-diseases> [accessed 2021-02-02]
2. Chronic diseases in Europe. European Chronic Disease Alliance. 2019. URL: <https://alliancechronicdiseases.org/wp-content/uploads/Guide-for-European-Commissioner-for-Health-Chronic-diseases-in-Europe-ECDA-2019.pdf> [accessed 2021-02-02]
3. Wang Y, Wang J. Modelling and prediction of global non-communicable diseases. BMC Public Health 2020 Jun 01;20(1):822 [FREE Full text] [doi: [10.1186/s12889-020-08890-4](https://doi.org/10.1186/s12889-020-08890-4)] [Medline: [32487173](https://pubmed.ncbi.nlm.nih.gov/32487173/)]
4. WHO guidelines on physical activity and sedentary behaviour. World Health Organization. URL: <https://www.who.int/publications/i/item/9789240015128> [accessed 2021-02-02]
5. Global action plan on physical activity 2018-2030: more active people for a healthier world. World Health Organization. 2018. URL: <https://apps.who.int/iris/bitstream/handle/10665/272722/9789241514187-eng.pdf> [accessed 2021-02-02]
6. McGuire S. World Health Organization. comprehensive implementation plan on maternal, infant, and young child nutrition. Geneva, Switzerland, 2014. Adv Nutr 2015 Jan;6(1):134-135 [FREE Full text] [doi: [10.3945/an.114.007781](https://doi.org/10.3945/an.114.007781)] [Medline: [25593153](https://pubmed.ncbi.nlm.nih.gov/25593153/)]
7. Eng T. The eHealth Landscape: A Terrain Map of Emerging Information and Communication Technologies in Health and Health Care. Princeton, NJ: Robert Wood Johnson Foundation; 2001.

8. mHealth: new horizons for health through mobile technologies: second global survey on eHealth. World Health Organization. 2011. URL: <https://apps.who.int/iris/handle/10665/44607> [accessed 2021-09-26]
9. Wang Y, Xue H, Huang Y, Huang L, Zhang D. A systematic review of application and effectiveness of mhealth interventions for obesity and diabetes treatment and self-management. *Adv Nutr* 2017 May 15;8(3):449-462 [FREE Full text] [doi: [10.3945/an.116.014100](https://doi.org/10.3945/an.116.014100)] [Medline: [28507010](https://pubmed.ncbi.nlm.nih.gov/28507010/)]
10. Batch BC, Spratt SE, Blalock DV, Benditz C, Weiss A, Dolor RJ, et al. General behavioral engagement and changes in clinical and cognitive outcomes of patients with type 2 diabetes using the time2focus mobile app for diabetes education: pilot evaluation. *J Med Internet Res* 2021 Jan 20;23(1):e17537 [FREE Full text] [doi: [10.2196/17537](https://doi.org/10.2196/17537)] [Medline: [33470947](https://pubmed.ncbi.nlm.nih.gov/33470947/)]
11. Hamilton SJ, Mills B, Birch EM, Thompson SC. Smartphones in the secondary prevention of cardiovascular disease: a systematic review. *BMC Cardiovasc Disord* 2018 Feb 07;18(1):25 [FREE Full text] [doi: [10.1186/s12872-018-0764-x](https://doi.org/10.1186/s12872-018-0764-x)] [Medline: [29415680](https://pubmed.ncbi.nlm.nih.gov/29415680/)]
12. Müller AM, Alley S, Schoeppe S, Vandelanotte C. *Int J Behav Nutr Phys Act* 2016 Oct 10;13(1):109 [FREE Full text] [doi: [10.1186/s12966-016-0434-2](https://doi.org/10.1186/s12966-016-0434-2)] [Medline: [27724911](https://pubmed.ncbi.nlm.nih.gov/27724911/)]
13. Haque MS, Kangas M, Jämsä T. A persuasive mhealth behavioral change intervention for promoting physical activity in the workplace: feasibility randomized controlled trial. *JMIR Form Res* 2020 May 04;4(5):e15083 [FREE Full text] [doi: [10.2196/15083](https://doi.org/10.2196/15083)] [Medline: [32364506](https://pubmed.ncbi.nlm.nih.gov/32364506/)]
14. McCarroll R, Eyles H, Ni MC. Effectiveness of mobile health (mHealth) interventions for promoting healthy eating in adults: a systematic review. *Prev Med* 2017 Dec;105:156-168. [doi: [10.1016/j.ypmed.2017.08.022](https://doi.org/10.1016/j.ypmed.2017.08.022)] [Medline: [28882743](https://pubmed.ncbi.nlm.nih.gov/28882743/)]
15. Hakala S, Rintala A, Immonen J, Karvanen J, Heinonen A, Sjögren T. Effectiveness of technology-based distance interventions promoting physical activity: systematic review, meta-analysis and meta-regression. *J Rehabil Med* 2017 Jan 31;49(2):97-105 [FREE Full text] [doi: [10.2340/16501977-2195](https://doi.org/10.2340/16501977-2195)] [Medline: [28112356](https://pubmed.ncbi.nlm.nih.gov/28112356/)]
16. Motivating patients to use smartphone health apps. Cision PR Web. 2011. URL: <http://www.prweb.com/releases/2011/04/prweb5268884.htm> [accessed 2021-02-02]
17. Lentferink AJ, Oldenhuis HK, de GM, Polstra L, Velthuisen H, van GJE. Key components in ehealth interventions combining self-tracking and persuasive ecoaching to promote a healthier lifestyle: a scoping review. *J Med Internet Res* 2017 Aug 01;19(8):e277 [FREE Full text] [doi: [10.2196/jmir.7288](https://doi.org/10.2196/jmir.7288)] [Medline: [28765103](https://pubmed.ncbi.nlm.nih.gov/28765103/)]
18. Kelders SM, Kok RN, Ossebaard HC, Van GJWC. Persuasive system design does matter: a systematic review of adherence to web-based interventions. *J Med Internet Res* 2012 Nov 14;14(6):e152 [FREE Full text] [doi: [10.2196/jmir.2104](https://doi.org/10.2196/jmir.2104)] [Medline: [23151820](https://pubmed.ncbi.nlm.nih.gov/23151820/)]
19. Vo V, Auroy L, Sarradon-Eck A. Patients' perceptions of mhealth apps: meta-ethnographic review of qualitative studies. *JMIR Mhealth Uhealth* 2019 Jul 10;7(7):e13817 [FREE Full text] [doi: [10.2196/13817](https://doi.org/10.2196/13817)] [Medline: [31293246](https://pubmed.ncbi.nlm.nih.gov/31293246/)]
20. Asbjørnsen RA, Smedsrød ML, Solberg Nes L, Wentzel J, Varsi C, Hjelmæsæth J, et al. Persuasive system design principles and behavior change techniques to stimulate motivation and adherence in electronic health interventions to support weight loss maintenance: scoping review. *J Med Internet Res* 2019 Jun 21;21(6):e14265 [FREE Full text] [doi: [10.2196/14265](https://doi.org/10.2196/14265)] [Medline: [31228174](https://pubmed.ncbi.nlm.nih.gov/31228174/)]
21. Popay J, Roberts H, Sowden A, Petticrew M, Arai L, Rodgers M, et al. Guidance on the conduct of narrative synthesis in systematic reviews. Lancaster University. 2006. URL: <https://www.lancaster.ac.uk/media/lancaster-university/content-assets/documents/fhm/dhr/chir/NSsynthesisguidanceVersion1-April2006.pdf> [accessed 2021-02-02]
22. Moher D, Liberati A, Tetzlaff J, Altman DG. Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. *PLoS Med* 2009 Jul 21;6(7):e1000097 [FREE Full text] [doi: [10.1371/journal.pmed.1000097](https://doi.org/10.1371/journal.pmed.1000097)] [Medline: [19621072](https://pubmed.ncbi.nlm.nih.gov/19621072/)]
23. Page MJ, McKenzie JE, Bossuyt PM, Boutron I, Hoffmann TC, Mulrow CD, et al. The PRISMA 2020 statement: an updated guideline for reporting systematic reviews. *BMJ* 2021 Mar 29;372:n71 [FREE Full text] [doi: [10.1136/bmj.n71](https://doi.org/10.1136/bmj.n71)] [Medline: [33782057](https://pubmed.ncbi.nlm.nih.gov/33782057/)]
24. Istepanian R, Alanzi T. Mobile health (m-health): evidence-based progress or scientific retrogression. In: *Biomedical Information Technology* (2nd ed). Dammam: Academic Press; 2020:717-733.
25. Pollock A, Berge E. How to do a systematic review. *Int J Stroke* 2018 Feb 17;13(2):138-156. [doi: [10.1177/1747493017743796](https://doi.org/10.1177/1747493017743796)] [Medline: [29148960](https://pubmed.ncbi.nlm.nih.gov/29148960/)]
26. Thomas J, Harden A. Methods for the thematic synthesis of qualitative research in systematic reviews. *BMC Med Res Methodol* 2008;8:45 [FREE Full text] [doi: [10.1186/1471-2288-8-45](https://doi.org/10.1186/1471-2288-8-45)] [Medline: [18616818](https://pubmed.ncbi.nlm.nih.gov/18616818/)]
27. Evaluation of methods in health care and intervention in social services: a method book. The Swedish National Agency for Medical and Social Evaluation. URL: <https://www.sbu.se/sv/metod/sbus-metodbok/> [accessed 2021-04-07]
28. Forster A, Buykx P, Martin N, Sadler S, Southgate B, Rockcliffe L, et al. Using affective judgement to increase physical activity in British adults. *Health Promot Int* 2018 Aug 01;33(4):648-656 [FREE Full text] [doi: [10.1093/heapro/dax004](https://doi.org/10.1093/heapro/dax004)] [Medline: [28334769](https://pubmed.ncbi.nlm.nih.gov/28334769/)]
29. Groarke JM, Richmond J, Mc Sharry J, Groarke A, Harney OM, Kelly MG, et al. Acceptability of a mobile health behavior change intervention for cancer survivors with obesity or overweight: nested mixed methods study within a randomized controlled trial. *JMIR Mhealth Uhealth* 2021 Feb 16;9(2):e18288 [FREE Full text] [doi: [10.2196/18288](https://doi.org/10.2196/18288)] [Medline: [33591290](https://pubmed.ncbi.nlm.nih.gov/33591290/)]

30. Mummah SA, King AC, Gardner CD, Sutton S. Iterative development of Vegethon: a theory-based mobile app intervention to increase vegetable consumption. *Int J Behav Nutr Phys Act* 2016 Aug 08;13:90 [FREE Full text] [doi: [10.1186/s12966-016-0400-z](https://doi.org/10.1186/s12966-016-0400-z)] [Medline: [27501724](https://pubmed.ncbi.nlm.nih.gov/27501724/)]
31. Paul L, Brewster S, Wyke S, McFadyen AK, Sattar N, Gill JM, et al. Increasing physical activity in older adults using STARFISH, an interactive smartphone application (app); a pilot study. *J Rehabil Assist Technol Eng* 2017 Apr 05;4:2055668317696236 [FREE Full text] [doi: [10.1177/2055668317696236](https://doi.org/10.1177/2055668317696236)] [Medline: [31186925](https://pubmed.ncbi.nlm.nih.gov/31186925/)]
32. Roberts AL, Potts HW, Koutoukidis DA, Smith L, Fisher A. Breast, prostate, and colorectal cancer survivors' experiences of using publicly available physical activity mobile apps: qualitative study. *JMIR Mhealth Uhealth* 2019 Jan 04;7(1):e10918 [FREE Full text] [doi: [10.2196/10918](https://doi.org/10.2196/10918)]
33. Whitelock V, Kersbergen I, Higgs S, Aveyard P, Halford JC, Robinson E. User experiences of a smartphone-based attentive eating app and their association with diet and weight loss outcomes: thematic and exploratory analyses from a randomized controlled trial. *JMIR Mhealth Uhealth* 2020 Oct 02;8(10):e16780 [FREE Full text] [doi: [10.2196/16780](https://doi.org/10.2196/16780)] [Medline: [33006564](https://pubmed.ncbi.nlm.nih.gov/33006564/)]
34. Baretta D, Perski O, Steca P. Exploring users' experiences of the uptake and adoption of physical activity apps: longitudinal qualitative study. *JMIR Mhealth Uhealth* 2019 Feb 08;7(2):e11636 [FREE Full text] [doi: [10.2196/11636](https://doi.org/10.2196/11636)] [Medline: [30735143](https://pubmed.ncbi.nlm.nih.gov/30735143/)]
35. Degroote L, Van Dyck D, De Bourdeaudhuij I, De Paepe A, Crombez G. Acceptability and feasibility of the mHealth intervention 'MyDayPlan' to increase physical activity in a general adult population. *BMC Public Health* 2020 Jun 29;20(1):1032 [FREE Full text] [doi: [10.1186/s12889-020-09148-9](https://doi.org/10.1186/s12889-020-09148-9)] [Medline: [32600352](https://pubmed.ncbi.nlm.nih.gov/32600352/)]
36. Poppe L, Van der Mispel C, Crombez G, De Bourdeaudhuij I, Schroé H, Verloigne M. How users experience and use an ehealth intervention based on self-regulation: mixed-methods study. *J Med Internet Res* 2018 Oct 01;20(10):e10412 [FREE Full text] [doi: [10.2196/10412](https://doi.org/10.2196/10412)] [Medline: [30274961](https://pubmed.ncbi.nlm.nih.gov/30274961/)]
37. Karppinen P, Oinas-Kukkonen H, Alahäivälä T, Jokelainen T, Keränen A, Salonurmi T, et al. Persuasive user experiences of a health behavior change support system: a 12-month study for prevention of metabolic syndrome. *Int J Med Inform* 2016 Dec;96:51-61. [doi: [10.1016/j.ijmedinf.2016.02.005](https://doi.org/10.1016/j.ijmedinf.2016.02.005)] [Medline: [26992482](https://pubmed.ncbi.nlm.nih.gov/26992482/)]
38. Nurmi J, Knittle K, Ginchev T, Khattak F, Helf C, Zwickl P, et al. Engaging users in the behavior change process with digitalized motivational interviewing and gamification: development and feasibility testing of the precious app. *JMIR Mhealth Uhealth* 2020 Jan 30;8(1):e12884 [FREE Full text] [doi: [10.2196/12884](https://doi.org/10.2196/12884)] [Medline: [32003750](https://pubmed.ncbi.nlm.nih.gov/32003750/)]
39. Ribeiro N, Moreira L, Almeida AMP, Santos-Silva F. Pilot study of a smartphone-based intervention to promote cancer prevention behaviours. *Int J Med Inform* 2017 Dec;108:125-133. [doi: [10.1016/j.ijmedinf.2017.10.013](https://doi.org/10.1016/j.ijmedinf.2017.10.013)] [Medline: [29132617](https://pubmed.ncbi.nlm.nih.gov/29132617/)]
40. Lindqvist A, Rutberg S, Söderström E, Ek A, Alexandrou C, Maddison R, et al. User perception of a smartphone app to promote physical activity through active transportation: inductive qualitative content analysis within the smart city active mobile phone intervention (SCAMPI) study. *JMIR Mhealth Uhealth* 2020 Aug 05;8(8):e19380 [FREE Full text] [doi: [10.2196/19380](https://doi.org/10.2196/19380)] [Medline: [32755889](https://pubmed.ncbi.nlm.nih.gov/32755889/)]
41. Wichmann F, Pischke CR, Jürgens D, Darmann-Finck I, Koppelin F, Lippke S, et al. Requirements for (web-based) physical activity interventions targeting adults above the age of 65 years - qualitative results regarding acceptance and needs of participants and non-participants. *BMC Public Health* 2020 Jun 11;20(1):907 [FREE Full text] [doi: [10.1186/s12889-020-08927-8](https://doi.org/10.1186/s12889-020-08927-8)] [Medline: [32527251](https://pubmed.ncbi.nlm.nih.gov/32527251/)]
42. Ceasar JN, Claudel SE, Andrews MR, Tamura K, Mitchell V, Brooks AT, et al. Community engagement in the development of an mhealth-enabled physical activity and cardiovascular health intervention (Step It Up): pilot focus group study. *JMIR Form Res* 2019 Jan 04;3(1):e10944 [FREE Full text] [doi: [10.2196/10944](https://doi.org/10.2196/10944)] [Medline: [30684422](https://pubmed.ncbi.nlm.nih.gov/30684422/)]
43. Du Y, Dennis B, Rhodes SL, Sia M, Ko J, Jiwani R, et al. Technology-assisted self-monitoring of lifestyle behaviors and health indicators in diabetes: qualitative study. *JMIR Diabetes* 2020 Aug 28;5(3):e21183 [FREE Full text] [doi: [10.2196/21183](https://doi.org/10.2196/21183)] [Medline: [32857056](https://pubmed.ncbi.nlm.nih.gov/32857056/)]
44. Miropolsky EM, Scott Baker K, Abbey-Lambertz M, Syrjala K, Chow EJ, Ceballos R, et al. Participant perceptions on a Fitbit and Facebook intervention for young adult cancer survivors: a qualitative study. *J Adolesc Young Adult Oncol* 2020 Jun 01;9(3):410-417 [FREE Full text] [doi: [10.1089/jayao.2019.0072](https://doi.org/10.1089/jayao.2019.0072)] [Medline: [31928489](https://pubmed.ncbi.nlm.nih.gov/31928489/)]
45. Joseph RP, Keller C, Vega-López S, Adams MA, English R, Hollingshead K, et al. A culturally relevant smartphone-delivered physical activity intervention for African American women: development and initial usability tests of smart walk. *JMIR Mhealth Uhealth* 2020 Mar 02;8(3):e15346 [FREE Full text] [doi: [10.2196/15346](https://doi.org/10.2196/15346)] [Medline: [32130198](https://pubmed.ncbi.nlm.nih.gov/32130198/)]
46. Vaghefi I, Tulu B. The continued use of mobile health apps: insights from a longitudinal study. *JMIR Mhealth Uhealth* 2019 Aug 29;7(8):e12983 [FREE Full text] [doi: [10.2196/12983](https://doi.org/10.2196/12983)] [Medline: [31469081](https://pubmed.ncbi.nlm.nih.gov/31469081/)]
47. Alturki R, Gay V. The development of an arabic weight-loss app Akser Waznk: qualitative results. *JMIR Form Res* 2019 Mar 14;3(1):e11785 [FREE Full text] [doi: [10.2196/11785](https://doi.org/10.2196/11785)] [Medline: [30869639](https://pubmed.ncbi.nlm.nih.gov/30869639/)]
48. Gilson ND, Pavey TG, Wright OR, Vandelanotte C, Duncan MJ, Gomersall S, et al. The impact of an m-Health financial incentives program on the physical activity and diet of Australian truck drivers. *BMC Public Health* 2017 May 18;17(1):467 [FREE Full text] [doi: [10.1186/s12889-017-4380-y](https://doi.org/10.1186/s12889-017-4380-y)] [Medline: [28521767](https://pubmed.ncbi.nlm.nih.gov/28521767/)]

49. Mauch CE, Laws RA, Prichard I, Maeder AJ, Wycherley TP, Golley RK. Commercially available apps to support healthy family meals: user testing of app utility, acceptability, and engagement. *JMIR Mhealth Uhealth* 2021 May 07;9(5):e22990 [FREE Full text] [doi: [10.2196/22990](https://doi.org/10.2196/22990)] [Medline: [33960951](https://pubmed.ncbi.nlm.nih.gov/33960951/)]
50. Lieffers JRL, Valaitis RF, George T, Wilson M, Macdonald J, Hanning RM. A qualitative evaluation of the eaTracker mobile app. *Nutrients* 2018 Oct 09;10(10):1462 [FREE Full text] [doi: [10.3390/nu10101462](https://doi.org/10.3390/nu10101462)] [Medline: [30304766](https://pubmed.ncbi.nlm.nih.gov/30304766/)]
51. Michie S, Richardson M, Johnston M, Abraham C, Francis J, Hardeman W, et al. The behavior change technique taxonomy (v1) of 93 hierarchically clustered techniques: building an international consensus for the reporting of behavior change interventions. *Ann Behav Med* 2013 Aug;46(1):81-95. [doi: [10.1007/s12160-013-9486-6](https://doi.org/10.1007/s12160-013-9486-6)] [Medline: [23512568](https://pubmed.ncbi.nlm.nih.gov/23512568/)]
52. Dennison L, Morrison L, Conway G, Yardley L. Opportunities and challenges for smartphone applications in supporting health behavior change: qualitative study. *J Med Internet Res* 2013 Apr;15(4):e86 [FREE Full text] [doi: [10.2196/jmir.2583](https://doi.org/10.2196/jmir.2583)] [Medline: [23598614](https://pubmed.ncbi.nlm.nih.gov/23598614/)]
53. Webb TL, Joseph J, Yardley L, Michie S. Using the internet to promote health behavior change: a systematic review and meta-analysis of the impact of theoretical basis, use of behavior change techniques, and mode of delivery on efficacy. *J Med Internet Res* 2010 Feb;12(1):e4 [FREE Full text] [doi: [10.2196/jmir.1376](https://doi.org/10.2196/jmir.1376)] [Medline: [20164043](https://pubmed.ncbi.nlm.nih.gov/20164043/)]
54. Kelders S. Understanding Adherence to Web-Based Interventions. Enschede, the Netherlands: University of Twente; 2012.
55. Shneiderman B, Plaisant C, Cohen M, Jacobs S. Designing the User Interface: Strategies for Effective Human-Computer Interaction. India: Pearson; 2010.
56. Broekhuizen K, Kroeze W, van PMNM, Oenema A, Brug J. A systematic review of randomized controlled trials on the effectiveness of computer-tailored physical activity and dietary behavior promotion programs: an update. *Ann Behav Med* 2012 Oct;44(2):259-286 [FREE Full text] [doi: [10.1007/s12160-012-9384-3](https://doi.org/10.1007/s12160-012-9384-3)] [Medline: [22767052](https://pubmed.ncbi.nlm.nih.gov/22767052/)]
57. Thomas Craig KJ, Morgan LC, Chen C, Michie S, Fusco N, Snowdon JL, et al. Systematic review of context-aware digital behavior change interventions to improve health. *Transl Behav Med* 2021 May 25;11(5):1037-1048. [doi: [10.1093/tbm/ibaa099](https://doi.org/10.1093/tbm/ibaa099)] [Medline: [33085767](https://pubmed.ncbi.nlm.nih.gov/33085767/)]
58. Strecher VJ, Seijts GH, Kok GJ, Latham GP, Glasgow R, DeVellis B, et al. Goal setting as a strategy for health behavior change. *Health Educ Q* 1995 May 04;22(2):190-200. [doi: [10.1177/109019819502200207](https://doi.org/10.1177/109019819502200207)] [Medline: [7622387](https://pubmed.ncbi.nlm.nih.gov/7622387/)]
59. Nikolaou CK, Lean MEJ. Mobile applications for obesity and weight management: current market characteristics. *Int J Obes (Lond)* 2017 Dec;41(1):200-202. [doi: [10.1038/ijo.2016.186](https://doi.org/10.1038/ijo.2016.186)] [Medline: [27780974](https://pubmed.ncbi.nlm.nih.gov/27780974/)]
60. Michie S, Abraham C, Whittington C, McAteer J, Gupta S. Effective techniques in healthy eating and physical activity interventions: a meta-regression. *Health Psychol* 2009 Nov;28(6):690-701. [doi: [10.1037/a0016136](https://doi.org/10.1037/a0016136)] [Medline: [19916637](https://pubmed.ncbi.nlm.nih.gov/19916637/)]
61. Cullen KW, Baranowski T, Smith SP. Using goal setting as a strategy for dietary behavior change. *J Am Diet Assoc* 2001 May;101(5):562-566. [doi: [10.1016/S0002-8223\(01\)00140-7](https://doi.org/10.1016/S0002-8223(01)00140-7)] [Medline: [11374350](https://pubmed.ncbi.nlm.nih.gov/11374350/)]
62. Johnson D, Deterding S, Kuhn K, Staneva A, Stoyanov S, Hides L. Gamification for health and wellbeing: a systematic review of the literature. *Internet Interv* 2016 Nov;6:89-106 [FREE Full text] [doi: [10.1016/j.invent.2016.10.002](https://doi.org/10.1016/j.invent.2016.10.002)] [Medline: [30135818](https://pubmed.ncbi.nlm.nih.gov/30135818/)]
63. White M, Dorman S. Receiving social support online: implications for health education. *Health Educ Res* 2001 Dec;16(6):693-707. [doi: [10.1093/her/16.6.693](https://doi.org/10.1093/her/16.6.693)] [Medline: [11780708](https://pubmed.ncbi.nlm.nih.gov/11780708/)]
64. Ribeiro N, Moreira L, Barros A, Almeida AM, Santos-Silva F. Guidelines for a cancer prevention smartphone application: a mixed-methods study. *Int J Med Inform* 2016 Oct;94:134-142. [doi: [10.1016/j.ijmedinf.2016.07.007](https://doi.org/10.1016/j.ijmedinf.2016.07.007)] [Medline: [27573321](https://pubmed.ncbi.nlm.nih.gov/27573321/)]
65. Peng W, Kanthawala S, Yuan S, Hussain SA. A qualitative study of user perceptions of mobile health apps. *BMC Public Health* 2016 Nov 14;16(1):1158 [FREE Full text] [doi: [10.1186/s12889-016-3808-0](https://doi.org/10.1186/s12889-016-3808-0)] [Medline: [27842533](https://pubmed.ncbi.nlm.nih.gov/27842533/)]

Abbreviations

mHealth: mobile health

Edited by A Kushniruk; submitted 14.10.21; peer-reviewed by S Stockwell, S Lalla-Edward, R Marshall; comments to author 20.01.22; revised version received 15.03.22; accepted 01.05.22; published 28.06.22.

Please cite as:

Bergevi J, Andermo S, Woldamanuel Y, Johansson UB, Hagströmer M, Rossen J
User Perceptions of eHealth and mHealth Services Promoting Physical Activity and Healthy Diets: Systematic Review
JMIR Hum Factors 2022;9(2):e34278
URL: <https://humanfactors.jmir.org/2022/2/e34278>
doi: [10.2196/34278](https://doi.org/10.2196/34278)
PMID: [35763339](https://pubmed.ncbi.nlm.nih.gov/35763339/)

©Julia Bergevi, Susanne Andermo, Yohannes Woldamanuel, Unn-Britt Johansson, Maria Hagströmer, Jenny Rossen. Originally published in JMIR Human Factors (<https://humanfactors.jmir.org>), 28.06.2022. This is an open-access article distributed under the terms of the Creative Commons Attribution License (<https://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Human Factors, is properly cited. The complete bibliographic information, a link to the original publication on <https://humanfactors.jmir.org>, as well as this copyright and license information must be included.

Original Paper

Assessing the Usability of a Clinical Decision Support System: Heuristic Evaluation

Hwayoung Cho¹, RN, PhD; Gail Keenan¹, RN, PhD; Olatunde O Madandola¹, RN, MPH; Fabiana Cristina Dos Santos¹, RN, MSN; Tamara G R Macieira¹, RN, PhD; Ragnhildur I Bjarnadottir¹, RN, MPH, PhD; Karen J B Priola¹, MSCIS; Karen Dunn Lopez², RN, PhD

¹College of Nursing, University of Florida, Gainesville, FL, United States

²College of Nursing, University of Iowa, Iowa City, IA, United States

Corresponding Author:

Hwayoung Cho, RN, PhD
College of Nursing
University of Florida
1225 Center Dr
Gainesville, FL, 32611
United States
Phone: 1 3522736347
Email: hcho@ufl.edu

Abstract

Background: Poor usability is a primary cause of unintended consequences related to the use of electronic health record (EHR) systems, which negatively impacts patient safety. Due to the cost and time needed to carry out iterative evaluations, many EHR components, such as clinical decision support systems (CDSSs), have not undergone rigorous usability testing prior to their deployment in clinical practice. Usability testing in the predeployment phase is crucial to eliminating usability issues and preventing costly fixes that will be needed if these issues are found after the system's implementation.

Objective: This study presents an example application of a systematic evaluation method that uses clinician experts with human-computer interaction (HCI) expertise to evaluate the usability of an electronic clinical decision support (CDS) intervention prior to its deployment in a randomized controlled trial.

Methods: We invited 6 HCI experts to participate in a heuristic evaluation of our CDS intervention. Each expert was asked to independently explore the intervention at least twice. After completing the assigned tasks using patient scenarios, each expert completed a *heuristic evaluation checklist* developed by Bright et al based on Nielsen's 10 heuristics. The experts also rated the overall severity of each identified heuristic violation on a scale of 0 to 4, where 0 indicates no problems and 4 indicates a usability catastrophe. Data from the experts' coded comments were synthesized, and the severity of each identified usability heuristic was analyzed.

Results: The 6 HCI experts included professionals from the fields of nursing (n=4), pharmaceutical science (n=1), and systems engineering (n=1). The mean overall severity scores of the identified heuristic violations ranged from 0.66 (*flexibility and efficiency of use*) to 2.00 (*user control and freedom* and *error prevention*), in which scores closer to 0 indicate a more usable system. The heuristic principle *user control and freedom* was identified as the most in need of refinement and, particularly by nonnursing HCI experts, considered as having major usability problems. In response to the heuristic *match between system and the real world*, the experts pointed to the reversed direction of our system's pain scale scores (1=severe pain) compared to those commonly used in clinical practice (typically 1=mild pain); although this was identified as a minor usability problem, its refinement was repeatedly emphasized by nursing HCI experts.

Conclusions: Our heuristic evaluation process is simple and systematic and can be used at multiple stages of system development to reduce the time and cost needed to establish the usability of a system before its widespread implementation. Furthermore, heuristic evaluations can help organizations develop transparent reporting protocols for usability, as required by Title IV of the 21st Century Cures Act. Testing of EHRs and CDSSs by clinicians with HCI expertise in heuristic evaluation processes has the potential to reduce the frequency of testing while increasing its quality, which may reduce clinicians' cognitive workload and errors and enhance the adoption of EHRs and CDSSs.

KEYWORDS

usability; heuristic; clinical decision support; electronic health record; expert review; evaluation; user interface; human-computer interaction

Introduction

Despite the great potential of electronic health records (EHRs), clinicians are often confronted with unintended consequences related to the use of these systems, which can negatively impact patient safety [1-3]. One of the primary reasons for these unforeseen challenges stems from the lack of or poorly executed usability testing of these systems [4-6].

Usability measures the quality of a user's experience when interacting with a system [7]. Recent evidence suggests that poor usability in EHRs is associated with an increase in clinicians' cognitive workload, EHR-related fatigue, burnout, work inefficiency, job dissatisfaction, and intentions to leave the job [8-10]. System acceptance and adoption are crucial and strongly associated with the usability of EHR systems [11-13]. To optimize the benefits of EHRs for clinicians and avoid any unintended consequences that adversely impact patient safety, it is imperative to establish a system's usability before its widespread implementation in real-world practice.

Usability evaluation methods are generally classified as expert- or user-based. Expert-based evaluations (eg, heuristic evaluations, cognitive walkthroughs, field observations) focus on ensuring that a system's functionality is optimized and evidence-based interface standards and norms are met [14,15]. Evidence-based interface standards have been developed by various researchers to answer the following questions: (1) Does the user interface conform to evidence-based design principles? (2) Can users accomplish a given task? (3) Are users satisfied with the way a system helps perform a task? and (4) Can users operate the system efficiently with a quality outcome? [16-19]. In contrast, user-based evaluations focus on a user's experience and interaction with a given system (eg, think-aloud method, interviews, focus groups, questionnaires) [14,15,20,21]. Although user-based usability testing shows differences in task performance between users who experienced difficulties and those who did not, expert-based usability testing focuses on "making things work" (ie, functionality) [12,14,20,22].

Clinical decision support systems (CDSSs) are specific components of EHRs that are frequently added and updated to reflect new evidence. CDSSs are defined as systems that provide clinicians with clinical knowledge and patient information that is "intelligently filtered and presented at appropriate times to improve patient care" [23]. When used as intended, CDSSs provide clinicians easy access to evidence-based information relevant to their decision-making process and can reduce their cognitive burden by minimizing the amount of information they must remember; these benefits enhance work efficiency, improve adherence to clinical guidelines, reduce the occurrence of medication errors, and prevent misdiagnoses [24-27]. Surprisingly, many CDSSs have not undergone rigorous usability and effectiveness testing prior to their deployment in

practice [28]. The testing of CDSSs' textual information and interfaces is critical in optimizing clinical decision-making and preventing errors in guidance.

A major challenge to establishing the usability of the CDSSs interfaced with EHRs has been the cost and time needed to carry out rigorous, iterative evaluations [21,29]. Attempting to fix usability issues after widespread deployment results in much higher costs than if done before implementation. Although usability studies should be iteratively conducted at multiple stages during system development [15], usability evaluations of health information technologies are often conducted during only a single stage of development [14]. In previous studies of CDSSs developed for clinicians that include nurses, usability testing was typically conducted either at an early stage for prototyping using an expert-based method [27,30] or after their deployment in practice using a user-based method [31-33]. Nurses participated in the evaluations mostly as a target user [31-33]; they may act as an expert—although they do not have usability expertise—after training by a usability expert to conduct the evaluation [27].

We believe that combining user- and expert-based evaluations has the potential to improve the efficiency and effectiveness of a system. In a user-based evaluation, the average cost per general user (ie, nonclinicians) is US \$171, and at least twenty users are needed to identify 95% of the usability problems in a system [34,35]. Conducting iterative usability evaluations of EHRs and CDSSs with clinician-users is even more costly and time-consuming because recruiting them in clinical studies remains challenging [36,37]. In an expert-based evaluation, 3 to 5 expert evaluators are recommended [38], and 3 experts can identify 80%-90% of the usability problems [39]. Although both types of evaluation are valuable in testing EHRs and CDSSs [27,30-33], the stage of development often dictates the choice of the usability evaluation conducted. However, the predeployment phase, which occurs after prototyping, is the most crucial phase since eliminating usability issues in this phase avoids the costly fixes that will be needed if they are found after a system's implementation [40,41]. Therefore, involving both experts and users in a late-stage (ie, predeployment stage after prototyping) usability evaluation would be optimal.

In this study, we offer an example application of our heuristic evaluation process, which provides a low-cost, time-effective, and expert-based method that includes potential users (ie, clinician usability experts) to evaluate the usability of CDSSs prior to their deployment in clinical practice.

Methods

Heuristic Evaluation

A heuristic evaluation is a usability-inspection method commonly used in the field of human-computer interaction

(HCI) [16,21,38,39]. The heuristic evaluation proposed by Nielsen is an assessment conducted by a small group of evaluators using an evidence-based set of design guidelines called heuristics [38,42]. Heuristic evaluators, who are generally experts in HCI, examine a user interface and the system design according to the evidence-based interface standards.

Example of Heuristic Evaluation Method

The example application of our approach involved the systematic evaluation of an electronic intervention containing clinical decision support (CDS) that was being prepared for deployment and testing by nurses in a national randomized controlled trial (RCT). Prior to nationwide deployment, we conducted a heuristic evaluation with HCI experts to identify any violations of usability principles in the CDS intervention.

We chose the heuristic evaluation process based on Nielsen's 10 heuristics [42] and used a *heuristic evaluation checklist* developed by Bright et al [43]. The checklist facilitated each expert's systematic inspection of the system's user interface by judging its compliance with each usability factor through yes-or-no questioning and rating its overall severity for each of Nielsen's 10 heuristics [42] on a scale of 0 (no problems) to 4 (usability catastrophe). Our heuristic evaluation process included specific HCI experts with nursing informatics expertise (referred to as "nursing HCI experts") and general HCI experts (referred

to as "nursing HCI experts") to capture the views of both usability experts and clinician-users of our CDS intervention.

CDS Intervention Under Evaluation

The main components of the CDS intervention evaluated in this paper were nursing diagnoses [44], nursing outcomes [45] with current and expected ratings, and nursing interventions [46]. Through an iterative design process with users (ie, nurses), our study team had previously developed and pretested a desktop prototype intervention designed to evaluate the effectiveness of 3 different electronic CDS intervention display formats: (1) text, (2) table, and (3) graph (see [Figure 1](#)) [47-50]. The CDS intervention contained evidence-based suggestions for improving palliative care delivered via a modular EHR care planning system (see [Figure 2](#)).

Subsequently, our team was funded by the National Institutes of Health to conduct a national, remotely administered RCT of the previously developed intervention. A desktop prototype in the 3 display formats ([Figure 1](#)) underwent iterative, user-centered-design usability studies with users (ie, user-based evaluations) [47-50]; however, a web-based application was needed to remotely test the CDS intervention with a national sample of 200 nurses. As small interface changes can impact the overall usability of an electronic CDS intervention, our team chose to conduct a second phase of usability testing using expert perspectives (ie, expert-based evaluations).

Figure 1. Three types of display formats (reproduced with permission from the HANDS Research Team). NANDA-I: NANDA International nursing diagnosis; NIC: nursing intervention classification; NOC: nursing outcome classification; POC: plan of care.

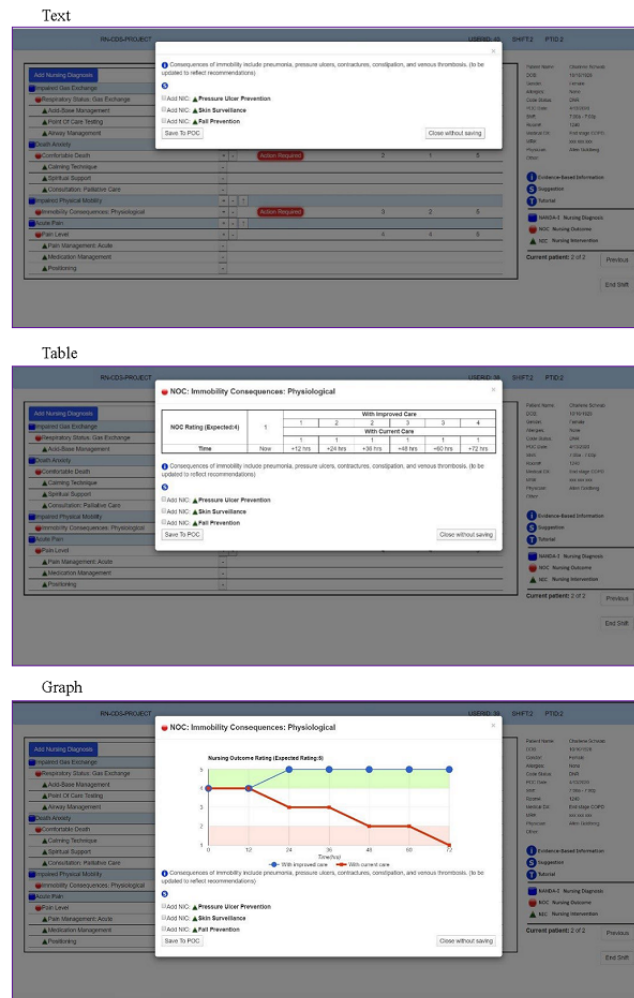
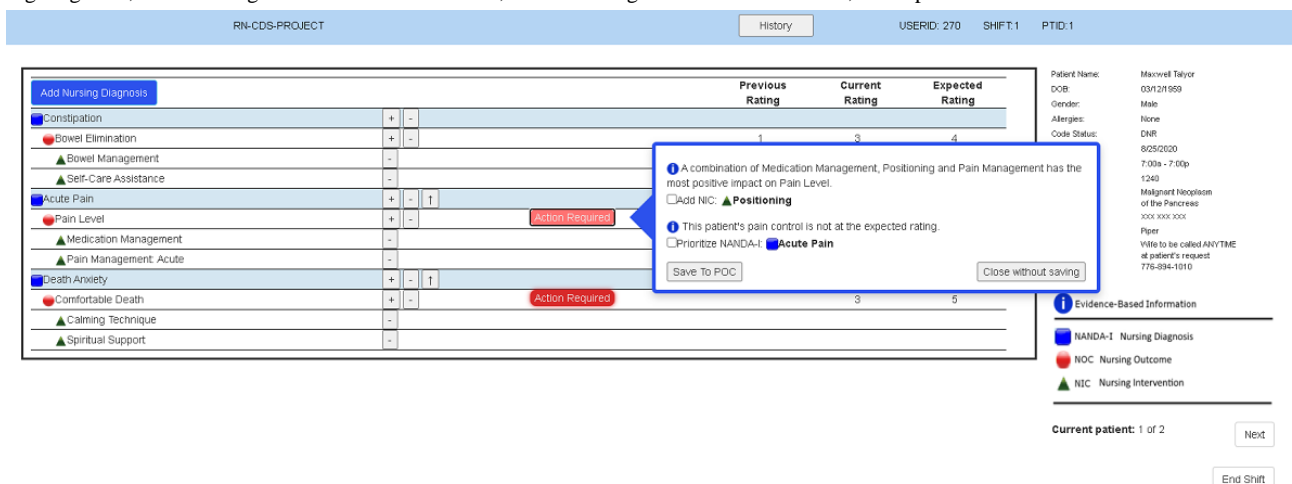


Figure 2. Clinical decision support suggestions (reproduced with permission from the HANDS Research Team). NANDA-I: NANDA International nursing diagnosis; NIC: nursing intervention classification; NOC: nursing outcome classification; POC: plan of care.



Sampling and Recruitment

We used purposive sampling to invite 6 HCI experts, including nursing and nonnursing HCI experts, to participate in this study from August 3, 2020, to September 11, 2020. The sample size was decided in accordance with current recommendations, which state that including more than 3 to 5 evaluators in a heuristic

evaluation is unlikely to yield additional useful information [38]. The main qualifications for participation were possession of a doctoral degree in the field of informatics and training in HCI. These qualifications were essential in this study since the quality of a heuristic evaluation is dependent on the skills and experience of the evaluators [22,51].

Procedure

Our heuristic evaluation was conducted virtually during the COVID-19 pandemic. Before the evaluation, each expert was given a standardized orientation using a Microsoft PowerPoint video and transcript about how the CDS intervention works. The experts were also presented with the 2 use cases shown in [Figure 3](#); these patient case scenarios require users (ie, nurses) to adjust their care plans to the unfolding clinical context. During the evaluation, each expert was asked to independently interact with the CDS intervention, ensuring unbiased evaluations from each evaluator. The experts were encouraged to explore the user interface of the entire CDS intervention at least twice.

After completing their given tasks using the use cases, each expert was asked to complete a *heuristic evaluation checklist* [42,43]. They were then asked to rate the overall severity of each identified heuristic violation on a scale of 0 to 4: 0 being no problems, 1 being a cosmetic problem only (ie, a fix can wait), 2 being a minor problem, 3 being a major problem, and 4 being a usability catastrophe (ie, requiring an immediate fix). Space was provided for the experts to add explanatory comments to identify the deficits of a usability factor and additional comments to justify each severity score. Since our upcoming clinical trial will test evidence-based suggestions using 3 information display formats (ie, text, table, and graph; see [Figure 1](#)), the *aesthetic and minimalist design* heuristic from the checklist was evaluated per display format.

Figure 3. Use cases describing patient scenarios. POC: plan of care.

Now that you have received report, you would typically round to the patients' rooms and do your initial assessment. Imagine these are your assessments and observations from your rounds.

Patient 1 Mr. Taylor:

- When you entered Mr. Taylor's room his eyes were filled with tears and he was gripping his rosary beads.
- When asked about his tears, he shared his fears about dying and about missing the many family events in the upcoming year (new grandchild, a wedding, and a graduation).
- Seeming embarrassed, he apologized for his tears and quickly wiped them away saying he needs to be strong for his family and they should not see him crying
- He reports his pain to be at a 2 on a 5 point scale with "1" being extreme pain and "5" being pain free.
- The nursing assistant reports that Mr. Taylor just had a large, formed bowel movement.

Patient 2 Mrs. Schwab:

- When you entered Mrs. Schwab's room her respirations were rapid but not labored (40-42/minute) when she was awake.
- She had no complaints of pain and appeared to doze off when left undisturbed.
- You notice a reddened area on her coccyx when you reposition Mrs. Schwab on her left side.
- While taking her vitals she became agitated and began to cry, expressing her fear of impending death and having "unfinished business"
- Mrs. Schwab's daughter is at the bedside and scolds her stating "Mom, don't talk like that, you aren't going anywhere."
- Her daughter also reports that Mrs. Schwab, is extremely weak, has difficulty bearing weight with 2 people helping her and absolutely is not to get out of bed.

Your task now is to enter the changes to both patients' POCs based on your assessments. To the extent possible, please assume the time constraints for documentation that you typically deal with in your practice.

HISTORY

Patient 1 Mr. Taylor's History: Our first patient is Mr. Taylor a 61 year old Caucasian male patient who was readmitted to the unit last evening with a diagnosis of end stage pancreatic cancer. He was diagnosed approximately 5 months ago and at that time was given a prognosis of 6 months. He was last admitted 2 weeks ago and during that admission elected to discontinue active cancer treatment and focus on acute symptom relief, with a desire to go home and spend the remaining time with his family. His code status is DNR. Up until yesterday, he has been cared for at home exclusively by family, his wife and three daughters. They have elected not to get hospice involved. About 24 hours prior to arrival in the emergency department the rating of his pain level was around a "1" on a 5 point scale (with 1 indicating extreme pain and 5 being no pain) so he was admitted to the floor.

His care is now focused on 2 NANDA-I diagnoses: Constipation & Acute Pain. His current NOC outcome rating for Bowel Elimination is 1, but the expected outcome should be a 4. His last bowel movement was 3 days ago, bowel sounds are diminished and his abdomen is slightly distended. I started him on a Bowel Management plan and gave him a suppository, but nothing has passed yet. He is unable to get out of bed, so he is using a bedpan; if his strength and mobility improves, he may be able to get up to a bedside commode. Mr. Taylor's current Pain Level NOC is a 2 and I expect that we should be able to get this under control soon and to a 5 with the IV morphine that was prescribed and by addressing his constipation issues. I instituted our unit Pain Management protocol, which should help as well.

Now that you have received report, you would typically round to the patient's rooms and do your initial assessment. Imagine these are your assessments and observations from your rounds.

- When you entered Mr. Taylor's room his eyes were filled with tears and he was gripping his rosary beads.
- When asked about his tears, he shared his fears about dying and about missing the many family events in the upcoming year (new grandchild, a wedding, and a graduation).
- Seeming embarrassed, he apologized for his tears and quickly wiped them away saying he needs to be strong for his family and they should not see him crying
- He reports his pain to be at a 2 on a 5 point scale with "1" being extreme pain and "5" being pain free.
- The nursing assistant reports that Mr. Taylor just had a large, formed bowel movement.

Your task now is to enter the changes to Mr. Taylor's POCs based on your assessments. To the extent possible, please assume the time constraints for documentation that you typically deal with in your practice.

Ethics Approval

The University of Florida Institutional Review Board reviewed and approved the addition of an evaluation of the intervention software by experts, with no subjects in the clinical trial involved (IRB201902611).

Data Analysis

Data analysis focused on the experts' comments and overall severity scores collected via the *heuristic evaluation checklist*. To capture the experts' perspectives on usability, we conducted deductive coding based on a pre-established set of guidelines (ie, heuristics). We developed a codebook for coding their comments using Microsoft Excel. Data from the coded comments were synthesized by 2 nursing informatics and HCI experts (HC and KDL), who were *not* participants in the heuristic evaluation, according to Nielsen's 10 usability

heuristics [38,42]. Differences in coding data were discussed until consensus was achieved.

Descriptive statistics were used to analyze the overall severity of the identified usability factors collected using the checklist. The mean and standard deviation of the overall severity score were calculated for each heuristic principle.

Results

The 6 HCI experts who participated in the heuristic evaluation were professionals in the fields of nursing (n=4), pharmaceutical sciences (n=1), and system engineering (n=1). The mean overall severity scores of the identified heuristic violations ranged from 0.66 (*flexibility and efficiency of use*) to 2.00 (*user control and freedom* and *error prevention*), in which scores closer to 0 indicate a more usable system. [Figure 4](#) depicts the mean

severity scores by heuristics and highlights the 4 highest scores. Table 1 organizes the evaluation’s mean severity scores and sample comments into Nielsen’s 10 usability heuristics.

The heuristic principles identified as the most in need of refinement were *user control and freedom* (mean 2.00, SD 1.09) and *error prevention* (mean 2.00, SD 1.09). Although all heuristics were identified as having major (ie, severity score of 3) and minor (ie, severity score of 2) usability problems, *user control and freedom* was considered a major usability issue particularly by nonnursing HCI experts, who pointed out that

users of the CDS intervention were unable to alter current and expected scores for nursing outcomes once the ratings were entered in. To improve this heuristic, the experts suggested that the “Undo” function should not be limited and to give users the ability to fix the entered scores. Similarly, after the “Action Required” menu was completed, it was no longer possible for users to select the “Undo” function to bring it up again. An example of this is shown in Figure 2, where the “Action Required” was choosing nursing interventions for the plan of care (POC) based on the decision support suggestions recommended by our CDS intervention.

Figure 4. Four highest mean severity scores by heuristic. Severity score from 0 to 4: no usability problems (0), cosmetic problem only (1), minor usability problem (2), major usability problem (3), and usability catastrophe (4).

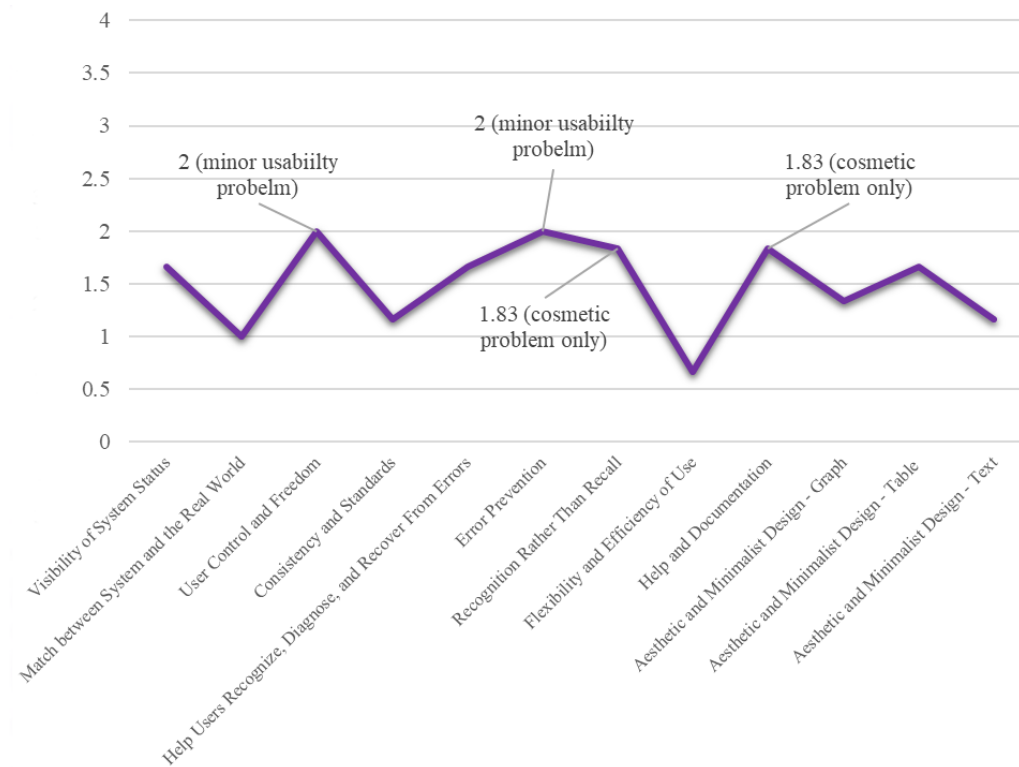


Table 1. Mean severity scores and sample comments from the heuristic evaluations.

Usability heuristic	Severity score ^a , mean (SD)	Sample comments
Visibility of system status	1.66 (1.21)	<ul style="list-style-type: none"> Unclear if care plan icons (circle, square, triangle) are clickable
Match between system and the real world	1.00 (1.09)	<ul style="list-style-type: none"> Pain scale in the CDS intervention, in which score 1 indicates “severe” pain is the opposite of common pain scales used in clinical practice (1 indicates “mild” pain)
User control and freedom	2.00 (1.09)	<ul style="list-style-type: none"> Limited “Undo” functionality
Consistency and standards	1.16 (1.16)	<ul style="list-style-type: none"> Unclear of formatting standards referred
Help users recognize, diagnose, and recover from errors	1.66 (1.63)	<ul style="list-style-type: none"> Error message is not informative as it doesn’t indicate where the error occurred
Error prevention	2.00 (1.09)	<ul style="list-style-type: none"> Need a warning message when clicking the minus button
Recognition rather than recall	1.83 (1.16)	<ul style="list-style-type: none"> Unclear what was undone
Flexibility and efficiency of use	0.66 (1.03)	<ul style="list-style-type: none"> Suggested helping users to find content on the site (hyperlinks, alphabetical index)
Help and documentation	1.83 (0.98)	<ul style="list-style-type: none"> Needs HELP function to inform on how the CDS intervention works
Aesthetic and minimalist design		
Graph format	1.33 (1.50)	<ul style="list-style-type: none"> Not visually appealing from similar blues/grey shades
Table format	1.66 (1.50)	<ul style="list-style-type: none"> Font is too small and difficult to read
Text format	1.16 (0.75)	<ul style="list-style-type: none"> No labels in the icons Suggested we use text section headers instead of icons

^aSeverity score from 0=best to 4=worst: no usability problems (0), cosmetic problem only (1), minor usability problem (2), major usability problem (3), and usability catastrophe (4).

In response to *error prevention*, the experts found the exit (x) button in the upper right corner of the “Action Required” menu to be confusing since 2 other options are also available: “Save To POC” and “Close without saving” in the lower left and right corners of the screen, respectively (Figure 5). To improve *error prevention*, the experts suggested that we provide the warning message shown in Figure 6 when the minus button is clicked; they also recommended that this warning message indicate where the error occurred to support the heuristic *help users recognize, diagnose, and recover from errors* (mean 1.66, SD 1.63).

The next heuristics identified as requiring the most improvement were *recognition rather than recall* (mean 1.83, SD 1.16) and *help and documentation* (mean 1.83, SD 0.98). *Recognition rather than recall* was considered a major usability problem

particularly by nonnursing HCI experts, who stated that clicking the “Undo” button to see what was undone should be recognizable to users. Regarding *help and documentation*, the experts emphasized the need for a “Help” or “Search” functionality to inform users of how our CDS intervention works (eg, how users can add a new nursing diagnosis) and reduce user errors when using the intervention.

Finally, for the heuristic *match between system and the real world* (mean 1.00, SD 1.09), the experts pointed to the reversed direction of our pain scale scores (1 indicating severe pain) compared to those commonly used in clinical practice (1 indicating mild pain; Figure 7). Although this usability issue was identified as minor, its refinement was repeatedly emphasized by nursing HCI experts.

Figure 5. Action Required menu (reproduced with permission from the HANDS Research Team). NIC: nursing intervention classification; NOC: nursing outcome classification; POC: plan of care.

The screenshot shows the RN-CDS-PROJECT interface. At the top, it displays 'RN-CDS-PROJECT', 'History', 'USERID: 271', 'SHIFT:1', and 'PTID:1'. Below this is a table of nursing diagnoses with columns for 'Previous Rating', 'Current Rating', and 'Expected Rating'. The 'Comfortable Death' diagnosis is highlighted with a red 'Action Required' label. A pop-up window titled 'NOC: Comfortable Death' is open, showing a table with 'NOC Rating (Expected:5)' and 'Time' columns. The table includes ratings for 'With Improved Care' and 'With Current Care' at various time intervals: Now, +12 hrs, +24 hrs, +36 hrs, +48 hrs, +60 hrs, and +72 hrs. The current rating is 3. Below the table, there is an information icon and text: 'Palliative care consultations help manage pain, symptoms, comorbidities, and patient/family communication. Add NIC: ▲ Consultation: Palliative Care'. Buttons for 'Save To POC' and 'Close without saving' are at the bottom of the pop-up.

Figure 6. Warning message (reproduced with permission from the HANDS Research Team). NANDA-I: NANDA International nursing diagnosis; NIC: nursing intervention classification; NOC: nursing outcome classification; POC: plan of care.

The screenshot shows a 'Warning' dialog box with a close button (X) in the top right corner. The text inside reads: 'All POC requirements are not fulfilled, make sure:'. Below this are three numbered items: '1. All new NOCs are rated.', '2. For every NANDA-I there is at least 1 NOC or delete it.', and '3. For every NOC there is at least 1 NIC or delete it.'. An 'Ok' button is located at the bottom left of the dialog box.

Figure 7. Pain scale scores (reproduced with permission from the HANDS Research Team). NANDA-I: NANDA International nursing diagnosis; NIC: nursing intervention classification; NOC: nursing outcome classification; POC: plan of care.

The screenshot shows the RN-CDS-PROJECT interface with a 'Rating the Nursing Outcome' dialog box open. The dialog box has a title bar and a close button (X). It contains two sections: 'The rating of patient now (at handoff)' and 'Expected rating:'. Each section has a 5-point scale with radio button options: '1 - Severe deviation from normal range', '2 - Substantial deviation from normal range', '3 - Moderate deviation from normal range', '4 - Mild deviation from normal range', and '5 - No deviation from normal range'. The 'Expected rating' section also includes the text 'The rating achievable by patient at discharge from your unit'. Below the scales are buttons for 'Save To POC' and 'Close without saving'. In the background, the nursing diagnosis table is visible, with 'Immobility Consequences: Physiological' highlighted in yellow and 'NR' in the 'Expected Rating' column. The right side of the interface shows patient information: 'Maxwell Taylor', 'DOB: 03/12/1959', 'Gender: Male', 'Allergies: None', 'Code Status: DNR', 'POC Date: 9/21/2020', 'Shift: 7:00a - 7:00p', 'Room#: 1240', 'Medical DX: Malignant Neoplasm of the Pancreas', 'MIR#: xxx xxxx', 'Physician: Piper', and 'Other: Wife to be called ANYTIME at patient's request 776-894-1010'. At the bottom right, it says 'Current patient: 1 of 2' with a 'Next' button and 'End Shift' button.

Discussion

Principal Findings

With the proliferation of EHRs and CDSSs in today's health care, rigorous and multistage usability evaluations are essential to the development of effective electronic systems; however, these evaluations are considered challenging due to the cost and time required to conduct them [21,29]. In this study, we provided an example application of a heuristic evaluation process that we used prior to the deployment of an electronic CDS intervention for our RCT study. The same process can be used with different EHRs and CDSSs and at multiple phases of development to provide high-quality, low-cost, and efficient usability assessments. This heuristic evaluation method can also help organizations develop transparent reporting on a system's usability, as required by Title IV of the 21st Century Cures Act [52]. As evidenced in this study, conducting this evaluation enabled us to detect unmet evidence-based usability principles of an electronic CDS intervention prior to its deployment.

This study took approximately 2 months (from August to September 2020) to locate and enlist the experts, distribute study materials, and compile the results. It is important to emphasize that this study was conducted during the global COVID-19 pandemic, which potentially affected the recruitment period as well as data collection. Thus, our process can likely be performed in a shorter period of time than the 2 months we experienced.

Through expert-based usability testing, we discovered major and minor usability problems in the user interface of an electronic CDS intervention prior to its deployment for use by users. Despite their benefits, heuristic evaluations are rarely reported for usability testing, especially in late-stage (ie, predeployment stage after prototyping) usability testing. Although user-based usability testing is effective in identifying major usability issues that affect user performance, a focus on user testing alone may lead to missed usability violations that users who do not have HCI expertise may not recognize [53-55]. Although unrecognized, these violations can decrease the system's usability, increase users' cognitive workload, create unintended consequences that threaten patient safety, and result in the EHR and CDSS being discontinued in practice. Future work should include a reevaluation of the CDS intervention after the recommendations against the heuristic violations have been implemented. In summary, heuristic evaluations have the potential to clarify usability issues within EHRs and CDSSs, not only after deployment but also before deployment, since they can be employed throughout various stages of system development [56]. Thus, this study reveals the value of including expert review methods at some point during the development process to ultimately achieve the goals of the system.

A heuristic evaluation with experts can identify minor usability problems that are often not detected in user testing but can be costly to fix if detected after a system's deployment [39]. Fixing usability problems after deployment or during maintenance

stages usually costs 40 to 100 times more than fixing them before deployment and in the development stage [40,41]; therefore, the early refinement of CDSSs using a heuristic evaluation process, such as the one described in this paper, ultimately reduces a system's overall development and redesign costs.

Since expert-based usability testing focuses on "making things work" in a natural and logical order, the experts in this study recommended changing the direction of our intervention's pain scale to range from 0 (no pain) to 4 (severe); this pain scale now matches those used in real-world clinical practice and would be intuitive to use. It is important to note that this usability problem was detected only by nursing HCI experts who have backgrounds in clinical nursing practice; this underscores the advantage of having a panel of experts who boasts skills and experience in the relevant clinical domains (eg, nursing, medicine), as well as in usability and HCI, when evaluating clinical technologies [51]. Our purposively selected panel of HCI experts, including nursing and nonnursing HCI experts, enabled us to identify significant usability problems that may have increased the likelihood of medical errors in real-world clinical settings, which is an important strength of this study.

Limitations

The limitations of this study were related to the experts' independent evaluations. To complete the evaluation, each expert used his or her own device (eg, desktop and laptop computers, tablets) with differing screen sizes; this could have influenced their evaluations of the CDS intervention. Nonetheless, to obtain an optimal idea of the intervention's general scope, we asked the experts to use Google Chrome's Incognito (ie, private) browser to access the intervention, as well as to carefully explore the user interface's screen layout and interaction structure at least twice [20].

Another potential limitation of our study is that we did not collect the demographic information of our study participants. We invited them to participate in our expert-based evaluation as HCI experts either with or without domain expertise.

Conclusions

Our heuristic evaluation process is simple, systematic, and theoretical and can ensure a system's optimal functionality. Beyond confirming that evidence-based interface standards and norms are met, our process can be used at multiple stages of system development before implementation (ie, predeployment phase after prototyping) to reduce the time and cost of the iterative evaluations needed to improve a system's usability after widespread implementation. A heuristic evaluation that includes HCI experts with domain expertise (ie, clinician HCI experts) has the potential to reduce the frequency of testing while increasing its quality, which may reduce clinicians' cognitive workload and EHR-related errors. Making this small investment in early refinement can reap sizable benefits to further enhance EHR and CDSS adoption and acceptance by various clinicians in real-world clinical practice.

Acknowledgments

This study was supported by the National Institute of Nursing Research, National Institutes of Health, under award number R01 NR018416-01. The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institute of Nursing Research, National Institutes of Health. We are grateful to Rishabh Garg for providing technical help.

Conflicts of Interest

None declared.

References

1. Harrison MI, Koppel R, Bar-Lev S. Unintended consequences of information technologies in health care--an interactive sociotechnical analysis. *J Am Med Inform Assoc* 2007;14(5):542-549 [FREE Full text] [doi: [10.1197/jamia.M2384](https://doi.org/10.1197/jamia.M2384)] [Medline: [17600093](https://pubmed.ncbi.nlm.nih.gov/17600093/)]
2. Reisman M. EHRs: the challenge of making electronic data usable and interoperable. *P T* 2017 Sep;42(9):572-575 [FREE Full text] [Medline: [28890644](https://pubmed.ncbi.nlm.nih.gov/28890644/)]
3. Malmberg ED, Phan TM, Harmon G, Nauert RF. Improving HIV/AIDS knowledge management using EHRs. *Online J Public Health Inform* 2012;4(3):ojphi.v4i3.4267 [FREE Full text] [doi: [10.5210/ojphi.v4i3.4267](https://doi.org/10.5210/ojphi.v4i3.4267)] [Medline: [23569643](https://pubmed.ncbi.nlm.nih.gov/23569643/)]
4. Yusof MM, Stergioulas L, Zugic J. Health information systems adoption: findings from a systematic review. *Stud Health Technol Inform* 2007;129(Pt 1):262-266. [Medline: [17911719](https://pubmed.ncbi.nlm.nih.gov/17911719/)]
5. Hettinger AZ, Melnick ER, Ratwani RM. Advancing electronic health record vendor usability maturity: progress and next steps. *J Am Med Inform Assoc* 2021 Apr 23;28(5):1029-1031 [FREE Full text] [doi: [10.1093/jamia/ocaa329](https://doi.org/10.1093/jamia/ocaa329)] [Medline: [33517394](https://pubmed.ncbi.nlm.nih.gov/33517394/)]
6. Howe JL, Adams KT, Hettinger AZ, Ratwani RM. Electronic health record usability issues and potential contribution to patient harm. *JAMA* 2018 Mar 27;319(12):1276-1278 [FREE Full text] [doi: [10.1001/jama.2018.1171](https://doi.org/10.1001/jama.2018.1171)] [Medline: [29584833](https://pubmed.ncbi.nlm.nih.gov/29584833/)]
7. Abran A, Khelifi A, Suryan W, Seffah A. Usability meanings and interpretations in ISO standards. *Software Quality Journal* 2003 Nov;11(4):325-338. [doi: [10.1023/A:1025869312943](https://doi.org/10.1023/A:1025869312943)]
8. Dunn Lopez K, Chin C, Leitão Azevedo RF, Kaushik V, Roy B, Schuh W, et al. Electronic health record usability and workload changes over time for provider and nursing staff following transition to new EHR. *Appl Ergon* 2021 May;93:103359. [doi: [10.1016/j.apergo.2021.103359](https://doi.org/10.1016/j.apergo.2021.103359)] [Medline: [33556884](https://pubmed.ncbi.nlm.nih.gov/33556884/)]
9. Kutney-Lee A, Brooks Carthon M, Sloane DM, Bowles KH, McHugh MD, Aiken LH. Electronic health record usability: associations with nurse and patient outcomes in hospitals. *Med Care* 2021 Jul 01;59(7):625-631. [doi: [10.1097/MLR.0000000000001536](https://doi.org/10.1097/MLR.0000000000001536)] [Medline: [33797506](https://pubmed.ncbi.nlm.nih.gov/33797506/)]
10. Khairat S, Coleman C, Ottmar P, Jayachander DI, Bice T, Carson SS. Association of electronic health record use with physician fatigue and efficiency. *JAMA Netw Open* 2020 Jun 01;3(6):e207385 [FREE Full text] [doi: [10.1001/jamanetworkopen.2020.7385](https://doi.org/10.1001/jamanetworkopen.2020.7385)] [Medline: [32515799](https://pubmed.ncbi.nlm.nih.gov/32515799/)]
11. Cho H, Porras T, Baik D, Beauchemin M, Schnell R. Understanding the predisposing, enabling, and reinforcing factors influencing the use of a mobile-based HIV management app: a real-world usability evaluation. *Int J Med Inform* 2018 Sep;117:88-95 [FREE Full text] [doi: [10.1016/j.ijmedinf.2018.06.007](https://doi.org/10.1016/j.ijmedinf.2018.06.007)] [Medline: [30032969](https://pubmed.ncbi.nlm.nih.gov/30032969/)]
12. Cho H, Porras T, Flynn G, Schnell R. Usability of a consumer health informatics tool following completion of a clinical trial: focus group study. *J Med Internet Res* 2020 Jun 15;22(6):e17708 [FREE Full text] [doi: [10.2196/17708](https://doi.org/10.2196/17708)] [Medline: [32538796](https://pubmed.ncbi.nlm.nih.gov/32538796/)]
13. Davis FD. Perceived usefulness, perceived ease of use, and user acceptance of information technology. *MIS Quarterly* 1989 Sep;13(3):319-340 [FREE Full text] [doi: [10.2307/249008](https://doi.org/10.2307/249008)]
14. Yen P, Bakken S. Review of health information technology usability study methodologies. *J Am Med Inform Assoc* 2012 Jun;19(3):413-422 [FREE Full text] [doi: [10.1136/amiajnl-2010-000020](https://doi.org/10.1136/amiajnl-2010-000020)] [Medline: [21828224](https://pubmed.ncbi.nlm.nih.gov/21828224/)]
15. Cho H, Yen PY, Dowding D, Merrill JA, Schnell R. A multi-level usability evaluation of mobile health applications: a case study. *J Biomed Inform* 2018 Oct;86:79-89 [FREE Full text] [doi: [10.1016/j.jbi.2018.08.012](https://doi.org/10.1016/j.jbi.2018.08.012)] [Medline: [30145317](https://pubmed.ncbi.nlm.nih.gov/30145317/)]
16. Nielsen J. *Usability Engineering*. Boston, Massachusetts: Academic Press; 1993.
17. Norman DA. *The Design of Everyday Things*. New York: Currency Doubleday; 1990.
18. Nielsen J. 10 usability heuristics for user interface design. Nielsen Norman Group. 2020 Nov 15. URL: <https://www.nngroup.com/articles/ten-usability-heuristics/> [accessed 2022-04-29]
19. Shneiderman B. *Designing the User Interface*. Boston, Massachusetts: Addison Wesley; 1997.
20. Jaspers MWM. A comparison of usability methods for testing interactive health technologies: methodological aspects and empirical evidence. *Int J Med Inform* 2009 May;78(5):340-353. [doi: [10.1016/j.ijmedinf.2008.10.002](https://doi.org/10.1016/j.ijmedinf.2008.10.002)] [Medline: [19046928](https://pubmed.ncbi.nlm.nih.gov/19046928/)]
21. Holzinger A. Usability engineering methods for software developers. *Commun ACM* 2005 Jan;48(1):71-74 [FREE Full text] [doi: [10.1145/1039539.1039541](https://doi.org/10.1145/1039539.1039541)]

22. Cho H, Powell D, Pichon A, Thai J, Bruce J, Kuhns LM, et al. A mobile health intervention for HIV prevention among racially and ethnically diverse young men: usability evaluation. *JMIR mHealth uHealth* 2018 Sep 07;6(9):e11450 [[FREE Full text](#)] [doi: [10.2196/11450](https://doi.org/10.2196/11450)] [Medline: [30194060](https://pubmed.ncbi.nlm.nih.gov/30194060/)]
23. Teich JM, Osheroff JA, Pifer EA, Sittig DF, Jenders RA, CDS Expert Review Panel. Clinical decision support in electronic prescribing: recommendations and an action plan: report of the joint clinical decision support workgroup. *J Am Med Inform Assoc* 2005 ;12(4):365-376 [[FREE Full text](#)] [doi: [10.1197/jamia.M1822](https://doi.org/10.1197/jamia.M1822)] [Medline: [15802474](https://pubmed.ncbi.nlm.nih.gov/15802474/)]
24. Muhiyaddin R, Abd-Alrazaq AA, Househ M, Alam T, Shah Z. The impact of Clinical Decision Support Systems (CDSS) on physicians: a scoping review. *Stud Health Technol Inform* 2020 Jun 26;272:470-473. [doi: [10.3233/SHTI200597](https://doi.org/10.3233/SHTI200597)] [Medline: [32604704](https://pubmed.ncbi.nlm.nih.gov/32604704/)]
25. Jia P, Zhang L, Chen J, Zhao P, Zhang M. The effects of clinical decision support systems on medication safety: an overview. *PLoS One* 2016 Dec 15;11(12):e0167683 [[FREE Full text](#)] [doi: [10.1371/journal.pone.0167683](https://doi.org/10.1371/journal.pone.0167683)] [Medline: [27977697](https://pubmed.ncbi.nlm.nih.gov/27977697/)]
26. Sutton RT, Pincock D, Baumgart DC, Sadowski DC, Fedorak RN, Kroeker KI. An overview of clinical decision support systems: benefits, risks, and strategies for success. *NPJ Digit Med* 2020;3:17 [[FREE Full text](#)] [doi: [10.1038/s41746-020-0221-y](https://doi.org/10.1038/s41746-020-0221-y)] [Medline: [32047862](https://pubmed.ncbi.nlm.nih.gov/32047862/)]
27. Yuan MJ, Finley GM, Long J, Mills C, Johnson RK. Evaluation of user interface and workflow design of a bedside nursing clinical decision support system. *Interact J Med Res* 2013 Jan 31;2(1):e4 [[FREE Full text](#)] [doi: [10.2196/ijmr.2402](https://doi.org/10.2196/ijmr.2402)] [Medline: [23612350](https://pubmed.ncbi.nlm.nih.gov/23612350/)]
28. Dunn Lopez K, Gephart SM, Raszewski R, Sousa V, Shehorn LE, Abraham J. Integrative review of clinical decision support for registered nurses in acute care settings. *J Am Med Inform Assoc* 2017 Mar 01;24(2):441-450 [[FREE Full text](#)] [doi: [10.1093/jamia/ocw084](https://doi.org/10.1093/jamia/ocw084)] [Medline: [27330074](https://pubmed.ncbi.nlm.nih.gov/27330074/)]
29. Schnall R, Cho H, Liu J. Health Information Technology Usability Evaluation Scale (Health-ITUES) for usability assessment of mobile health technology: validation study. *JMIR mHealth uHealth* 2018 Jan 05;6(1):e4 [[FREE Full text](#)] [doi: [10.2196/mhealth.8851](https://doi.org/10.2196/mhealth.8851)] [Medline: [29305343](https://pubmed.ncbi.nlm.nih.gov/29305343/)]
30. Reeder B, Drake C, Ozkaynak M, Wald HL. Usability testing of a mobile clinical decision support app for urinary tract infection diagnosis in nursing homes. *J Gerontol Nurs* 2019 Jul 01;45(7):11-17. [doi: [10.3928/00989134-20190408-01](https://doi.org/10.3928/00989134-20190408-01)] [Medline: [30985907](https://pubmed.ncbi.nlm.nih.gov/30985907/)]
31. Muhindo M, Bress J, Kalanda R, Armas J, Danziger E, Kanya MR, et al. Implementation of a newborn clinical decision support software (NoviGuide) in a rural district hospital in eastern Uganda: feasibility and acceptability study. *JMIR mHealth uHealth* 2021 Feb 19;9(2):e23737 [[FREE Full text](#)] [doi: [10.2196/23737](https://doi.org/10.2196/23737)] [Medline: [33605886](https://pubmed.ncbi.nlm.nih.gov/33605886/)]
32. Marcolino MS, Oliveira JAQ, Cimini CCR, Maia JX, Pinto VSOA, Sá TQV, et al. Development and implementation of a decision support system to improve control of hypertension and diabetes in a resource-constrained area in Brazil: mixed methods study. *J Med Internet Res* 2021 Jan 11;23(1):e18872 [[FREE Full text](#)] [doi: [10.2196/18872](https://doi.org/10.2196/18872)] [Medline: [33427686](https://pubmed.ncbi.nlm.nih.gov/33427686/)]
33. Koskela T, Sandström S, Mäkinen J, Liira H. User perspectives on an electronic decision-support tool performing comprehensive medication reviews - a focus group study with physicians and nurses. *BMC Med Inform Decis Mak* 2016 Jan 22;16:6 [[FREE Full text](#)] [doi: [10.1186/s12911-016-0245-z](https://doi.org/10.1186/s12911-016-0245-z)] [Medline: [26801630](https://pubmed.ncbi.nlm.nih.gov/26801630/)]
34. Sova DH, Nielsen J. 234 tips and tricks for recruiting users as participants in usability studies. Nielson Norman Group. URL: https://media.nngroup.com/media/reports/free/How_To_Recruit_Participants_for_Usability_Studies.pdf [accessed 2022-04-29]
35. Faulkner L. Beyond the five-user assumption: benefits of increased sample sizes in usability testing. *Behav Res Methods Instrum Comput* 2003 Aug;35(3):379-383. [doi: [10.3758/bf03195514](https://doi.org/10.3758/bf03195514)] [Medline: [14587545](https://pubmed.ncbi.nlm.nih.gov/14587545/)]
36. Broyles LM, Rodriguez KL, Price PA, Bayliss NK, Sevick MA. Overcoming barriers to the recruitment of nurses as participants in health care research. *Qual Health Res* 2011 Dec;21(12):1705-1718. [doi: [10.1177/1049732311417727](https://doi.org/10.1177/1049732311417727)] [Medline: [21844286](https://pubmed.ncbi.nlm.nih.gov/21844286/)]
37. Asch S, Connor SE, Hamilton EG, Fox SA. Problems in recruiting community-based physicians for health services research. *J Gen Intern Med* 2000 Aug;15(8):591-599 [[FREE Full text](#)] [doi: [10.1046/j.1525-1497.2000.02329.x](https://doi.org/10.1046/j.1525-1497.2000.02329.x)] [Medline: [10940152](https://pubmed.ncbi.nlm.nih.gov/10940152/)]
38. Nielsen J, Molich R. Heuristic evaluation of user interfaces. 1990 Mar 01 Presented at: CHI '90: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems; 1990 Apr 1-5; Seattle, Washington p. 249-256. [doi: [10.1145/97243.97281](https://doi.org/10.1145/97243.97281)]
39. Nielsen J. Finding usability problems through heuristic evaluation. 1992 Jun 01 Presented at: CHI '92: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems; 1992 May 3-7; Monterey, California p. 373-380. [doi: [10.1145/142750.142834](https://doi.org/10.1145/142750.142834)]
40. Boehm BW. Software engineering economics. In: Broy M, Denert E, eds. *Pioneers and Their Contributions to Software Engineering*. Berlin, Heidelberg: Springer; 2001 Presented at: sd&m Conference on Software Pioneers; 2001 Jun 28-29; Bonn, Germany p. 99-150. [doi: [10.1007/978-3-642-48354-7](https://doi.org/10.1007/978-3-642-48354-7)]
41. Pressman RS. *Software Engineering: A Practitioner's Approach*. 3rd ed. New York: McGraw-Hill; 1992.
42. Nielsen J. Enhancing the explanatory power of usability heuristics. 1994 Apr 24 Presented at: CHI '94: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems; 1994 Apr 24-28; Boston, Massachusetts p. 152-158. [doi: [10.1145/191666.191729](https://doi.org/10.1145/191666.191729)]

43. Bright TJ, Bakken S, Johnson SB. Heuristic evaluation of eNote: an electronic notes system. *AMIA Annu Symp Proc* 2006;2006:864 [FREE Full text] [Medline: 17238484]
44. NANDA International. In: Herdman TH, Kamitsuru S, editors. *NANDA International Nursing Diagnoses: Definitions & Classification*, 2018-2020. 11th ed. New York, NY: Thieme; 2018.
45. Moorhead S, Johnson M, Maas ML, Swanson E, editors. *Nursing Outcomes Classification (NOC): Measurement of Health Outcomes*. 5th ed. St. Louis, Missouri: Elsevier/Mosby; 2013.
46. Butcher HK, Bulechek GM, Dochterman JMM, Wagner CM. *Nursing Interventions Classification (NIC)*. 7th ed. St. Louis, Missouri: Elsevier/Mosby; 2018.
47. Keenan GM, Lopez KD, Yao Y, Sousa VEC, Stifter J, Febretti A, et al. Toward meaningful care plan clinical decision support: feasibility and effects of a simulated pilot study. *Nurs Res* 2017;66(5):388-398 [FREE Full text] [doi: 10.1097/NNR.000000000000234] [Medline: 28858147]
48. Lopez KD, Febretti A, Stifter J, Johnson A, Wilkie DJ, Keenan G. Toward a more robust and efficient usability testing method of clinical decision support for nurses derived from nursing electronic health record data. *Int J Nurs Knowl* 2017 Oct;28(4):211-218 [FREE Full text] [doi: 10.1111/2047-3095.12146] [Medline: 27337939]
49. Lopez KD, Wilkie DJ, Yao Y, Sousa V, Febretti A, Stifter J, et al. Nurses' numeracy and graphical literacy: informing studies of clinical decision support interfaces. *J Nurs Care Qual* 2016;31(2):124-130 [FREE Full text] [doi: 10.1097/NCQ.000000000000149] [Medline: 26323050]
50. Stifter J, Sousa VEC, Febretti A, Dunn Lopez K, Johnson A, Yao Y, et al. Acceptability of clinical decision support interface prototypes for a nursing electronic health record to facilitate supportive care outcomes. *Int J Nurs Knowl* 2018 Oct;29(4):242-252 [FREE Full text] [doi: 10.1111/2047-3095.12178] [Medline: 28926204]
51. Po S, Howard S, Vetere F, Skov M. Heuristic evaluation and mobile usability: bridging the realism gap. Berlin, Germany: Springer; 2004 Presented at: *Mobile Human-Computer Interaction - MobileHCI 2004*; 2004 Sep 13-16; Glasgow, United Kingdom p. 49-60. [doi: 10.1007/978-3-540-28637-0_5]
52. 21st Century Cures Act, Public Law 114–255. 114th United States Congress. 2016 Dec 13. URL: <https://www.congress.gov/114/plaws/publ255/PLAW-114publ255.pdf> [accessed 2022-04-29]
53. Manhartsberger M, Zellhofer N. Eye tracking in usability research: what users really see. In: *Empowering software quality: how can usability engineering reach these goals?*. 2005 Presented at: *1st Usability Symposium*; 2005 Nov 8; Vienna, Austria p. 141-152 URL: <https://www.usability.at/ueberuns/EyetrackinginUsability.pdf>
54. Schiessl M, Duda S, Thölke A, Fischer R. Eye tracking and its application in usability and media research. *MMI-interaktiv Journal* 2003;6:41-50 [FREE Full text]
55. Cooke L, Cuddihy E. Using eye tracking to address limitations in think-aloud protocol. 2005 Presented at: *IPCC 2005. Proceedings. International Professional Communication Conference*, 2005; 2005 Jul 10-13; Limerick, Ireland p. 653-658. [doi: 10.1109/ipcc.2005.1494236]
56. Pertiwi AAP, Fraczkowski D, Stogis SL, Lopez KD. Using heuristic evaluation to improve sepsis alert usability. *Crit Care Nurs Clin North Am* 2018 Jun;30(2):297-309. [doi: 10.1016/j.cnc.2018.02.011] [Medline: 29724447]

Abbreviations

- CDS:** clinical decision support
- CDSS:** clinical decision support system
- EHR:** electronic health record
- HCI:** human-computer interaction
- POC:** plan of care
- RCT:** randomized controlled trial

Edited by A Kushniruk; submitted 02.07.21; peer-reviewed by F Yu, B Chaudhry, D Scherer, C Smith, R Marshall; comments to author 06.11.21; revised version received 31.01.22; accepted 19.04.22; published 10.05.22.

Please cite as:

Cho H, Keenan G, Madandola OO, Dos Santos FC, Macieira TGR, Bjarnadottir RI, Priola KJB, Dunn Lopez K
Assessing the Usability of a Clinical Decision Support System: Heuristic Evaluation
JMIR Hum Factors 2022;9(2):e31758
URL: <https://humanfactors.jmir.org/2022/2/e31758>
doi:10.2196/31758
PMID:35536613

©Hwayoung Cho, Gail Keenan, Olatunde O Madandola, Fabiana Cristina Dos Santos, Tamara G R Macieira, Ragnhildur I Bjarnadottir, Karen J B Priola, Karen Dunn Lopez. Originally published in JMIR Human Factors (<https://humanfactors.jmir.org>), 10.05.2022. This is an open-access article distributed under the terms of the Creative Commons Attribution License (<https://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Human Factors, is properly cited. The complete bibliographic information, a link to the original publication on <https://humanfactors.jmir.org>, as well as this copyright and license information must be included.

Original Paper

Adoption of a Postoperative Pain Self-Report Tool: Qualitative Study

Bram Thiel^{1*}, MPA, MSc; Inez Iao^{2*}, MSc; Joris Smid³, MSc; Emmy de Wit⁴, PhD; Seppe Koopman⁵, MD, PhD; Bart Geerts⁶, MD, MBA, PhD; Marc Godfried¹, MD, PhD; Cor Kalkman⁷, MD, PhD, Prof Dr

¹Department of Anesthesiology, OLVG Hospital, Amsterdam, Netherlands

²Department of Services and Solutions Delivery, Philips Benelux, Amsterdam, Netherlands

³Department of Cardiology, OLVG Hospital, Amsterdam, Netherlands

⁴Faculty of Science, Athena Institute, Vu University, Amsterdam, Netherlands

⁵Department of Anesthesiology, Maastad Hospital, Rotterdam, Netherlands

⁶Department of Intensive Care, Spaarne Gasthuis Hospital, Haarlem, Netherlands

⁷Department of Anesthesia and Intensive Care, University Medical Centre Utrecht, Utrecht, Netherlands

*these authors contributed equally

Corresponding Author:

Bram Thiel, MPA, MSc

Department of Anesthesiology

OLVG Hospital

Oosterpark 9

Amsterdam, 1091AC

Netherlands

Phone: 31 0205994773

Email: b.thiel@olvg.nl

Abstract

Background: With electronic technologies, patients are provided with tools to easily acquire information and to manage and record their own health status. eHealth interventions are already broadly applied to perioperative care. In a similar way, we aimed to utilize a smartphone application to enable postoperative patients to partially self-manage their postoperative pain. The results of a previously performed proof-of-concept study regarding the application were promising, and nurses as well as patients were optimistic regarding this innovative mobile application. Nevertheless, in reality, it appears that the usage and overall implementation of this application have stagnated since its introduction. Problems with innovation adoption are not novel; various studies have been conducted to explore the reasons for low implementation success of eHealth applications and indicated that adoption is influenced by multiple organizational factors. This study investigated the influence of these organizational factors on the adoption process, aiming to provide more insight in the dos and don'ts for implementing eHealth in the working processes of hospital care.

Objective: This study aimed to provide insight in how to successfully implement a technological eHealth innovation in a general nonacademic hospital.

Methods: A qualitative study was conducted to explore organizational factors affecting the innovation adoption process. Data were collected by conducting semistructured one-on-one interviews with 11 stakeholders. The data were analyzed using thematic analysis identifying overarching themes.

Results: Absorptive capacity, referred to as an organization's dynamic capability pertaining to knowledge creation and utilization that enhances an organization's ability to gain and sustain a competitive advantage, was regarded as the most influential factor on the application's adoption. Accordingly, it appeared that innovation adoption is mainly determined by the capability and willingness to assimilate and transform new information into productive use and the ability to absorb a novel innovation. Absorptive capacity was found to be influenced by the innovation's benefit and the sense of ownership and responsibility. Organizational readiness and management support were also regarded as essential since absorptive capacity seemed to be mediated by these factors. The size of the hospital influenced eHealth adoption by the amount of resources available and by its organizational structure.

Conclusions: In conclusion, absorptive capacity is essential for eHealth adoption, and it is mediated by management support and organizational readiness. It is recommended to increase the degree of willingness and ability to adopt an eHealth innovation by enhancing the relevance, engaging stakeholders, and assigning appropriate leaders to offer guidance.

(*JMIR Hum Factors* 2022;9(2):e33706) doi:[10.2196/33706](https://doi.org/10.2196/33706)

KEYWORDS

innovation; eHealth adoption model, mobile health; pain; self-report; perioperative medicine; postoperative pain; surgery

Introduction

To date, hospital care, admission potential, urgent and nonurgent care, and health care professionals are under pressure due to problems such as the aging population, the rising prevalence of chronic comorbid diseases, and overall increasing health care costs [1,2]. On top of these already existing problems came the COVID-19 pandemic. These growing demands accelerate the surge in eHealth innovations [3-5].

With electronic technologies, patients are provided with tools to easily acquire information and to manage and record their own health status. eHealth interventions are already broadly applied to perioperative care (eg, remote monitoring, educational websites, and telerehabilitation) [6,7]. In a similar way, we aimed to utilize an eHealth tool (the OLVG Pain App) with the objective to improve the efficiency and quality of postoperative pain management. This eHealth tool enables the patient to partially self-manage their postoperative pain, as they record their own postoperative pain intensity in their electronic medical record. The results of a previously performed proof-of-concept study regarding the OLVG Pain App were promising, and nurses as well as patients were optimistic regarding this innovative mobile application [8].

Nevertheless, in reality, it appears that the usage and overall implementation of this application have stagnated since its introduction. Problems with innovation adoption are not novel; various studies have been conducted to explore the reasons for low implementation success of eHealth applications and indicated that adoption is influenced by multiple organizational factors such as technological knowledge and skills, financial aspects, social and organizational support, and a lack of education and training [9-13]. Therefore, the aim of this study was to provide insight on how to successfully implement a technological eHealth innovation in a general nonacademic hospital. Accordingly, the research question was: “How can the

adoption process of the PainApp be understood with regards to the organizational factors within a general hospital?”

Methods

Study Design

A qualitative study was conducted between March 1, 2020, and July 31, 2020. Stakeholders involved with the development and implementation process of the application were interviewed to provide an in-depth understanding of how the context of a general hospital can facilitate or hamper the adoption of eHealth. The study was conducted and is reported according to the CONSORT-EHEALTH (Consolidated Standards of Reporting Trials of Electronic and Mobile Health Applications and onLine TeleHealth) checklist (V.1.6.1) [14], the consolidated criteria for reporting qualitative research (COREQ) [15], and the “Qualitative research: standards, challenges, methodological guidelines” by Malterud [16].

Recruitment

For this study, the perceptions of stakeholders active within the departments of Information Technology (IT), Electronic Medical Record (EMR), Anesthesiology, and Nursing Staff Convention from 2 hospitals were investigated. Both OLVG Hospital and Maasstad Hospital are general hospitals providing surgical care for 23,000 patients annually.

The sample size was determined by the concept of “information power,” which depends on the relevance of the participants included [17,18]. Based on this premise and similar studies, 11 participants were considered satisfactory, as the selected participants were highly informative and significant actors in the innovation procedure (Table 1, participant characteristics). Participants from specific departments within the OLVG Hospital and Maasstad Hospital were provided by personal contacts from the supervisors of this study.

Table 1. Participant characteristics.

Identification number	Hospital	Position	Age (years)	Gender	Work experience (years)	Technical background
P1	OLVG Hospital	Anesthetist	52	M ^a	17	6 years of postoperative home monitoring and digitalized preoperative screening
P2	OLVG Hospital	Nurse, application specialist	42	M	18	6 years as an application specialist for the EHR ^b
P3	Maastad Hospital	Anesthetist	39	M	6	1 year as a CMIO ^c
P4	OLVG Hospital	Neurologist	39	M	6	3 years as an innovation specialist, human-centered design, and eHealth implementation
P5 LKru	OLVG Hospital	Department manager	45	M	10	5 years as a manager of the EHR and patient participation
P6	OLVG Hospital	Nurse team leader	37	F ^d	15	3 years as a key user of the EPIC EHR
P7	OLVG Hospital	Department manager	- ^e	M	-	6 years as a manager of the ICT ^f department
P8	OLVG Hospital	Nurse team leader	-	F	28	10 years as a nurse team leader in the neurosurgery department
P9	OLVG Hospital	Operational manager	-	F	39	4 years as a manager of the neurosurgery business unit
P10	OLVG Hospital	Pulmonologist	62	M	30	2 years as a CMIO
P11	Maastad Hospital	Clinical computer scientist	34	M	5	Specialty connectivity between medical devices and the EHR

^aM: male.

^bEHR: electronic health record.

^cCMIO: Chief Medical Information Officer.

^dF: female.

^eParticipant did not respond to the extra questions for age and years of work experience.

^fICT: information communication and technology.

Data Collection

The data were collected during semistructured one-on-one interviews performed by a research student (IK). The interviews were structured by the following topic questions: size of the hospital, top management support, organizational readiness, and centralization in decision-making and absorptive capacity. These topics were based on pre-established themes derived from the eHealth Adoption Model (eHAM) [19]. The eHAM combines elements of the diffusion of innovation theory and the technology-organization-environment framework, since these form a theoretical base of innovation in various sectors [20].

The interviews were conducted in Dutch; the interview guide was originally formulated in Dutch, followed by a translation into English for the purpose of this report (Multimedia Appendix 1). Prior to the interview, participants received some general information concerning the research topic, and informed consent was requested from the interviewee to allow for audio recording

of the interview. The interview commenced with introductory questions, with which the general opinion on eHealth innovations of the participant was established. After this introductory phase, topic questions regarding the concepts of the eHAM model were asked to explore their experiences with and perspectives on the influences of these factors. Lastly, specific closing questions recapping the themes were asked to evaluate the importance of each organizational factor. Furthermore, interviews were conducted online through Skype, Zoom, or FaceTime due to COVID-19 measures.

Data Analysis

Data collection and analysis progression were discussed during regular meetings with the researchers BT and JS and the research student IK. A thematic framework approach was utilized in order to analyze the qualitative data [21,22]. First, the audio recordings of the interviews were transcribed verbatim in Dutch. As for the validity, the interpretations of the interview were sent to the respective participants to check whether the interviews

were well understood and in line with their perspectives. After familiarization with the data, the transcripts were coded with an English coding guide that was developed according to the eHAM model. The coding guide included the 5 main concepts regarding the organizational factors, which were further differentiated into subconcepts. After coding, overarching themes and patterns were identified and labelled within each concept.

Ethics Approval

This study was conducted as part of the “Closing the loop” project approved by the Advisory board for Science and Research (ACWO) OLVG Hospital on December 30, 2019, with registration number WO 19.167.

Results

Themes

The results were derived from stakeholder interviews and yielded a total of more than 40 themes that are related to the eHAM [19]. For a visualized overview of all identified and related themes, see [Multimedia Appendix 2](#). Overall, the stakeholders considered absorptive capacity, top management support, and organizational readiness as the most essential factors for the adoption of eHealth innovation. For interview data, see [Multimedia Appendix 3](#). Moreover, the remaining 2 factors (ie, size of the hospital and centralization in decision-making) were considered as generic influences and are therefore briefly discussed.

Absorptive Capacity

Absorptive capacity appeared to be of great importance. Absorptive capacity refers to an organization’s “dynamic capability pertaining to knowledge creation and utilization that enhances an organization’s ability to gain and sustain a competitive advantage” [23].

The majority of stakeholders stated that the individual’s willingness is an essential aspect that can affect the degree of absorptive capacity. Next to the willingness to absorb new information, the ability to do so is also essential for successful innovation adoption.

Altogether, the participants mentioned that spreading new knowledge is more effective when done repetitively face-to-face, than, for instance, digitally through newsletters or by email. Clinical lessons and pilot tests can also be used to effectively introduce innovations or increase the skills of employees.

Another important aspect mentioned to influence absorptive capacity is the culture within the hospital, as a culture that is more “open” and stimulating in accepting new information can enhance the adaptability and flexibility of the organization.

Furthermore, this willingness and ability to adopt certain innovations also appeared to depend on various themes, such as personal characteristics, the context, and whether a sense of ownership and responsibility is present with the individual. The influence of personal characteristics on absorptive capacity became evident as 9 of 11 participants noted that features such as age (or generation), affinity with technology, and being an

early or late adopter can affect whether an individual absorbs an innovation. Accordingly, a team leader stated:

When talking about absorptive capacity, it also depends a lot on the type of person. How much information can you assimilate? Are you theoretical or more practical? How old are you? Do you have an affinity for innovation and technology? [P6, team leader]

Additionally, 6 (P9, P1, P4, P5, P3, P6) participants felt that hospital employees are commonly not very capable of acquiring, assimilating, transforming, and connecting new information to existing knowledge for productive use. They believed that hospital employees are extremely programmed to conform to protocols, which limits the potential for adoption and particularly affects their ability to embed innovations into practiced daily routines. An essential remark made by all participants is that the context in which the innovation adoption takes place has a significant role in whether individuals are willing to absorb it. It became clear that the severity or urgency of the problem that the innovation addresses and responds to should be sufficiently experienced by those who need to use the innovation.

In addition, all participants mentioned that the innovation’s relevance is a key factor for successful innovation adoption and that this mainly depends on whether there is a substantial benefit to fulfill the need of the user. Another theme that emerged from 6 interviews was that the innovation should induce a sense of ownership in order for employees to truly adopt the eHealth innovation.

Top Management Support

Next to the absorptive capacity, the support of top management also appears to be of great influence, as managers are recognized to be responsible for delivering the required resources to facilitate innovation adoption. An anesthetist described the top management’s influence accordingly:

I think they certainly have an impact. They have to provide the money to be able to introduce eHealth tools such as the OLVG Pain app and support ICT and so on. Thus, if they say we won’t be doing it, then it won’t happen. [P3, anesthetist]

In addition, top management also seems to be responsible for arranging employees and project structures for realization. All participants agree that commitment from the top and a supportive vision of innovations should be present for successful eHealth adoption. More importantly, one of the team leaders stated that all management layers should be committed, in particular the team leaders, since they are practically active at the site of implementation:

You want it to be implemented in the workplace, so you should focus on there. The employees have no idea of what is happening in the top of the organization. If it is transmitted by a team leader, then they will believe and follow it. [P8, team leader]

Organizational Readiness

Most of the participants experienced that the actual IT infrastructure was adequately present to facilitate technological

developments. However, P6 did mention that the hospitals lag behind some current technologies due to limited financial resources and complex privacy concerns:

For example, we still do not have a blood pressure meter that automatically sends the data into the patient data management system (PDMS) via Wi-Fi. Nowadays, it must be possible to automatically Bluetooth or Wi-Fi everything to the system; why is it not possible for us to realize this? [P6, team leader]

Some participants supported this decision as they believed that hospitals are not supposed to compete with third-party app developers:

I don't think we'll be any better at it, because the hospital is good at treating patients and our core business is not building an app. [P5, department manager]

Financial Readiness

The majority of participants mentioned that they are aware that hospitals have relatively little financial resources. Furthermore, most of the participants recognized that the financial readiness for eHealth innovations mainly depends on the innovation's significance and potential for success. P6 described this accordingly:

Yes, if an innovation doctor says: "Yes we have to do this and I know it costs 100 thousand euros," the innovation committee can still object by saying that there is no money for that, because it has to be spent on new patient beds, as that is simply more important than that innovation. [P6, team leader]

Size of the Hospital

Most of the participants agreed that the size of the hospital's influence on eHealth implementation is associated with the available resources and the organizational structure. For instance, an IT employee mentioned that a larger hospital allows for more available resources, which in turn increases the organizational readiness for the adoption of eHealth innovations. On the other hand, a larger hospital appears to have more management layers and is therefore considered to be more bureaucratic in its structure. One of the participants (P5) stated correspondingly:

So yes, you have resources, but then again you are so big and bureaucratic that the speed of implementing innovations really slows down. [P5, department manager]

Centralization in Decision-making

All participants agreed that centralization is required for a structured overview to focus on hospital-wide interests and a fairer selection procedure of all projects. P5 described that a centralized group can positively influence eHealth adoption by providing guidance and resources:

That central group helps with getting a sharp picture of whether the innovation really matters and if that

is what we want, and then they'll help bridge the connection with those who can support, build, implement, and train it. [P5, department manager]

Discussion

Principal Findings

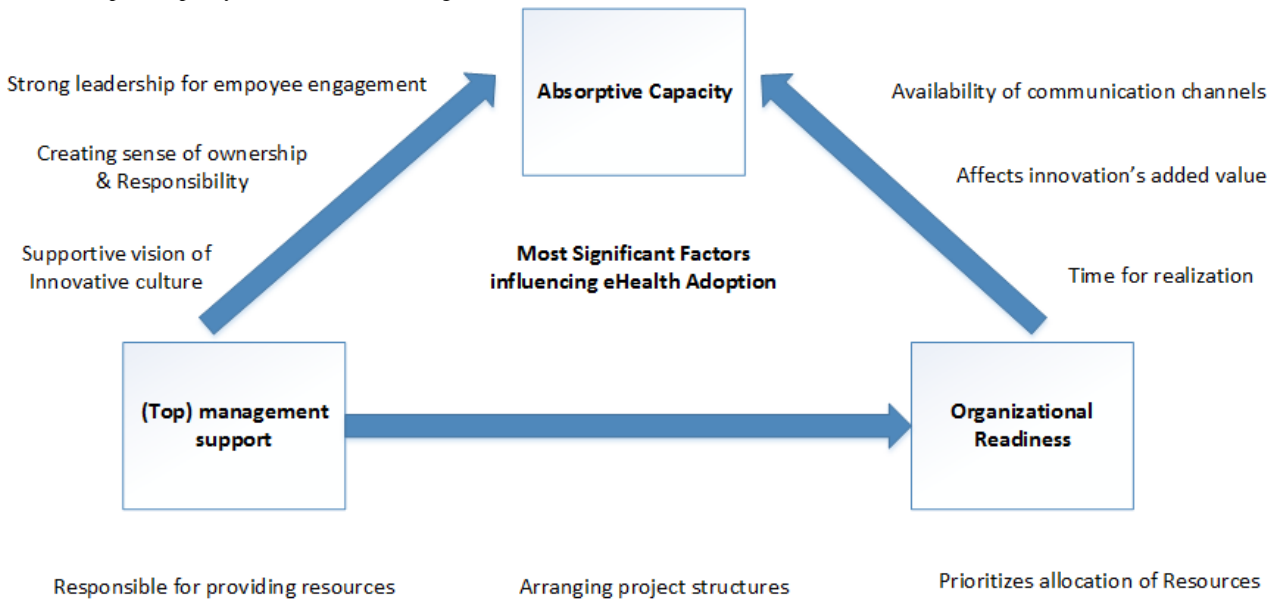
The results of this study demonstrate that absorptive capacity is regarded as the most important factor influencing eHealth adoption in a general hospital. In addition, the degree of absorptive capacity is predominantly mediated by the amount of management support and the organizational readiness (Figure 1). However, the size of the hospital and centralization in decision-making are rather generic influences on the innovation process. Previous studies show supportive findings regarding the importance of absorptive capacity in influencing eHealth adoption [24-27]. We found that "being attached to follow fixed guidelines" (ie, working by following strict protocols), "context" (ie, the way urgency and relevance of a specific issue push an innovation), and the absence of a sense of ownership and responsibility affected absorptive capacity. Being a "family owner" could contribute to the willingness for innovation and therefore could positively influence absorptive capacity [28]. Next to highlighting the importance of absorptive capacity itself, we also reveal that both management support and organizational readiness mediate this factor. Numerous other studies also present that absorptive capacity is influenced by organizational structure, culture, and communication [27,29-31].

Our findings regarding organizational readiness are in accordance with other studies suggesting that organizations with a favorable environment, the structure, and the required resources are more prone to absorb innovations [32,33].

For financial readiness, we found that a hospital's limited budget could be a barrier but would not be a determinant of the innovation's actual success. It seemed that, rather, the innovation's importance and potential to truly improve the quality of care would eventually determine the availability of financial resources.

The size of the hospital was found to influence the innovation process, as a larger size is positively associated with a greater amount of organizational resources, which in turn can facilitate better innovation implementation. This is in line with previous literature indicating that a greater hospital size influences the likelihood of successful adoption [34]. In contrast to this, our findings suggest that a greater size could also lead to a more bureaucratic structure and therefore even hamper innovation adoption despite its possession of more resources.

As for the centralization of decision-making, it became clear that this organizational factor differed in influence depending on the specific stage of the innovation process. Despite the potential positive influence of an acknowledgeable centralized group in guiding the innovation implementation, the stakeholders in our study ranked this factor as one of the least influential for successful adoption.

Figure 1. Absorptive capacity in relation to mediating factors.

Limitations

This study has several limitations; despite the recruitment of stakeholders from 2 different hospitals, the results and key messages of this study are mainly valid for the OLVG specifically, as the great majority of interviewees came from this hospital. This might have limited the acquirement of in-depth details on certain factors. The use of a conceptual model enabled this study to provide more insights into how the organizational factors influence adoption and to structure an overview of how these different factors relate to each other in terms of prioritization. However, the inclusion of such an extensive number of organizational factors may also have hampered a deeper understanding and exploration of each factor's influence on its own. Accordingly, this is also the reason why the hospital's "communication structures" have only been addressed briefly in our study and hence, lacks thoroughness in the findings.

Recommendations for Practice

The results revealed that adoption was prominently lacking on the innovation's benefit and on the sense of ownership and responsibility, which in turn negatively affected the absorptive capacity. Therefore, we suggest 3 focal points for policy.

The first focal point follows the "Quadruple Aim criteria" as a concept that focuses on examining the innovation according to the following 4 aims: lower cost of care, improved patient care, better health outcomes, and improved staff experience [35].

The second focal point increases the absorptive capacity from various aspects by using the value-sensitive design approach to engage all stakeholders from the beginning, creating a sense of ownership and responsibility [36].

The third focal point relates to our finding that management has a major influence on organizational culture and thus, also on the absorptive capacity. Accordingly, appropriate leaders who are truly able to stimulate an innovation-friendly learning culture should engage stakeholders. Therefore, the allocation of ambassadors or managers can support in overcoming resistance to change concerning stakeholder engagement and creating the innovation's benefit.

Conclusions

This study provides insight in how to successfully implement an eHealth innovation in a general hospital. The most important factor influencing eHealth adoption was absorptive capacity, which was mainly determined by the innovation's urgency and relevance, and a sense of ownership and responsibility. Additionally, we revealed that absorptive capacity is mediated by management support and organizational readiness. Three focal points for successful eHealth adoption are enhancing the innovation's relevance, adequately engaging stakeholders from the start, and allocating ambassadors or managers to support stakeholder engagement and to offer proper guidance and training.

Acknowledgments

This study was established from the internship in Master Management, Policy-analysis and entrepreneurship in Health and Life Science of II at OLVG Hospital, Course code: AM_1120. The internship was conducted as part of the Closing the Loop study approved by the local medical ethics committee and board of OLVG (WO 19.167). Researchers II and BT contributed equally to the final version of this manuscript.

Conflicts of Interest

None declared.

Multimedia Appendix 1

Interview guide.

[[DOCX File, 21 KB - humanfactors_v9i2e33706_app1.docx](#)]

Multimedia Appendix 2

Overview of the identified themes.

[[PNG File, 120 KB - humanfactors_v9i2e33706_app2.png](#)]

Multimedia Appendix 3

Interview data.

[[XLSX File \(Microsoft Excel File\), 95 KB - humanfactors_v9i2e33706_app3.xlsx](#)]

References

1. Hajat C, Stein E. The global burden of multiple chronic conditions: A narrative review. *Prev Med Rep* 2018 Dec;12:284-293 [[FREE Full text](#)] [doi: [10.1016/j.pmedr.2018.10.008](https://doi.org/10.1016/j.pmedr.2018.10.008)] [Medline: [30406006](https://pubmed.ncbi.nlm.nih.gov/30406006/)]
2. Harris G. World Health Organization launches new app to improve care for ageing population. *British Journal of Healthcare Management* 2020 Jan 02;26(1):48-49. [doi: [10.12968/bjhc.2019.0115](https://doi.org/10.12968/bjhc.2019.0115)]
3. Duplaga M. The acceptance of e-health solutions among patients with chronic respiratory conditions. *Telemed J E Health* 2013 Sep;19(9):683-691 [[FREE Full text](#)] [doi: [10.1089/tmj.2012.0306](https://doi.org/10.1089/tmj.2012.0306)] [Medline: [23734700](https://pubmed.ncbi.nlm.nih.gov/23734700/)]
4. Fairbrother P, Ure J, Hanley J, McCloughan L, Denvir M, Sheikh A, Telescot programme team. Telemonitoring for chronic heart failure: the views of patients and healthcare professionals - a qualitative study. *J Clin Nurs* 2014 Jan;23(1-2):132-144. [doi: [10.1111/jocn.12137](https://doi.org/10.1111/jocn.12137)] [Medline: [23451899](https://pubmed.ncbi.nlm.nih.gov/23451899/)]
5. Bakker M, Sloots J, Van Ommeren C, Kleberger A, Tabak M, Effing T, et al. Adherence in patients with COPD and heart failure using an eHealth self-management intervention. *European Respiratory Journal* 2019;54(suppl 63):PA744. [doi: [10.1183/13993003.congress-2019.PA744](https://doi.org/10.1183/13993003.congress-2019.PA744)]
6. van der Meij E, Anema JR, Otten RHJ, Huirne JAF, Schaafsma FG. The effect of perioperative e-Health interventions on the postoperative course: a systematic review of randomised and non-randomised controlled trials. *PLoS One* 2016 Jul 6;11(7):e0158612 [[FREE Full text](#)] [doi: [10.1371/journal.pone.0158612](https://doi.org/10.1371/journal.pone.0158612)] [Medline: [27383239](https://pubmed.ncbi.nlm.nih.gov/27383239/)]
7. Breteler M, Numan L, Ruurda J, van Hillegersberg R, van der Horst S, Dohmen D, et al. Wireless remote home monitoring of vital signs in patients discharged early after esophagectomy: observational feasibility study. *JMIR Perioper Med* 2020 Dec 04;3(2):e21705 [[FREE Full text](#)] [doi: [10.2196/21705](https://doi.org/10.2196/21705)] [Medline: [33393923](https://pubmed.ncbi.nlm.nih.gov/33393923/)]
8. Thiel B, Godfried M, van Huizen EC, Mooijer B, de Boer BA, van Mierlo RAAM, et al. Patient reported postoperative pain with a smartphone application: A proof of concept. *PLoS One* 2020;15(5):e0232082 [[FREE Full text](#)] [doi: [10.1371/journal.pone.0232082](https://doi.org/10.1371/journal.pone.0232082)] [Medline: [32384103](https://pubmed.ncbi.nlm.nih.gov/32384103/)]
9. Postema TRF, Peeters JM, Friele RD. Key factors influencing the implementation success of a home telecare application. *Int J Med Inform* 2012 Jun;81(6):415-423 [[FREE Full text](#)] [doi: [10.1016/j.ijmedinf.2011.12.003](https://doi.org/10.1016/j.ijmedinf.2011.12.003)] [Medline: [22226925](https://pubmed.ncbi.nlm.nih.gov/22226925/)]
10. Uscher-Pines L, Kahn JM. Barriers and facilitators to pediatric emergency telemedicine in the United States. *Telemed J E Health* 2014 Nov;20(11):990-996 [[FREE Full text](#)] [doi: [10.1089/tmj.2014.0015](https://doi.org/10.1089/tmj.2014.0015)] [Medline: [25238565](https://pubmed.ncbi.nlm.nih.gov/25238565/)]
11. Duplaga M. Searching for a role of nursing personnel in developing landscape of eHealth: factors determining attitudes toward key patient empowering applications. *PLoS One* 2016 Apr 6;11(4):e0153173 [[FREE Full text](#)] [doi: [10.1371/journal.pone.0153173](https://doi.org/10.1371/journal.pone.0153173)] [Medline: [27049525](https://pubmed.ncbi.nlm.nih.gov/27049525/)]
12. Valenzuela Espinoza A, De Smedt A, Guldolf K, Vandervorst F, Van Hooff R, Fernandez Tellez H, et al. Opinions and beliefs about telemedicine for emergency treatment during ambulance transportation and for chronic care at home. *Interact J Med Res* 2016 Mar 30;5(1):e9 [[FREE Full text](#)] [doi: [10.2196/ijmr.5015](https://doi.org/10.2196/ijmr.5015)] [Medline: [27029999](https://pubmed.ncbi.nlm.nih.gov/27029999/)]
13. Ariens LF, Schussler-Raymakers FM, Frima C, Flinterman A, Hamminga E, Arents BW, et al. Barriers and facilitators to eHealth use in daily practice: perspectives of patients and professionals in dermatology. *J Med Internet Res* 2017 Sep 05;19(9):e300 [[FREE Full text](#)] [doi: [10.2196/jmir.7512](https://doi.org/10.2196/jmir.7512)] [Medline: [28874336](https://pubmed.ncbi.nlm.nih.gov/28874336/)]
14. Eysenbach G, CONSORT-EHEALTH Group. CONSORT-EHEALTH: improving and standardizing evaluation reports of Web-based and mobile health interventions. *J Med Internet Res* 2011 Dec 31;13(4):e126 [[FREE Full text](#)] [doi: [10.2196/jmir.1923](https://doi.org/10.2196/jmir.1923)] [Medline: [22209829](https://pubmed.ncbi.nlm.nih.gov/22209829/)]
15. Tong A, Sainsbury P, Craig J. Consolidated criteria for reporting qualitative research (COREQ): a 32-item checklist for interviews and focus groups. *Int J Qual Health Care* 2007 Dec;19(6):349-357. [doi: [10.1093/intqhc/mzm042](https://doi.org/10.1093/intqhc/mzm042)] [Medline: [17872937](https://pubmed.ncbi.nlm.nih.gov/17872937/)]
16. Malterud K. Qualitative research: standards, challenges, and guidelines. *The Lancet* 2001 Aug;358(9280):483-488. [doi: [10.1016/s0140-6736\(01\)05627-6](https://doi.org/10.1016/s0140-6736(01)05627-6)]

17. Malterud K, Siersma VD, Guassora AD. Sample size in qualitative interview studies: guided by information power. *Qual Health Res* 2016 Nov;26(13):1753-1760. [doi: [10.1177/1049732315617444](https://doi.org/10.1177/1049732315617444)] [Medline: [26613970](https://pubmed.ncbi.nlm.nih.gov/26613970/)]
18. Guest G, Bunce A, Johnson L. How many interviews are enough? *Field Methods* 2016 Jul 21;18(1):59-82. [doi: [10.1177/1525822X05279903](https://doi.org/10.1177/1525822X05279903)]
19. Faber S, van Geenhuizen M, de Reuver M. eHealth adoption factors in medical hospitals: A focus on the Netherlands. *Int J Med Inform* 2017 Apr;100:77-89. [doi: [10.1016/j.ijmedinf.2017.01.009](https://doi.org/10.1016/j.ijmedinf.2017.01.009)] [Medline: [28241940](https://pubmed.ncbi.nlm.nih.gov/28241940/)]
20. Oliveira T, Martins MF. Literature review of information technology adoption models at firm level. *The Electronic Journal Information Systems Evaluation* 2011;14(1):110-121 [[FREE Full text](#)]
21. Pope C, Ziebland S, Mays N. Qualitative research in health care. Analysing qualitative data. *BMJ* 2000 Jan 08;320(7227):114-116 [[FREE Full text](#)] [doi: [10.1136/bmj.320.7227.114](https://doi.org/10.1136/bmj.320.7227.114)] [Medline: [10625273](https://pubmed.ncbi.nlm.nih.gov/10625273/)]
22. Gale NK, Heath G, Cameron E, Rashid S, Redwood S. Using the framework method for the analysis of qualitative data in multi-disciplinary health research. *BMC Med Res Methodol* 2013 Sep 18;13:117 [[FREE Full text](#)] [doi: [10.1186/1471-2288-13-117](https://doi.org/10.1186/1471-2288-13-117)] [Medline: [24047204](https://pubmed.ncbi.nlm.nih.gov/24047204/)]
23. Zahra S, George G. Absorptive capacity: a review, reconceptualization, and extension. *The Academy of Management Review* 2002 Apr;27(2):185. [doi: [10.2307/4134351](https://doi.org/10.2307/4134351)]
24. Lee C, Shim JP. An exploratory study of radio frequency identification (RFID) adoption in the healthcare industry. *European Journal of Information Systems* 2017 Dec 19;16(6):712-724. [doi: [10.1057/palgrave.ejis.3000716](https://doi.org/10.1057/palgrave.ejis.3000716)]
25. Caccia-Bava M, Guimaraes T, Harrington SJ. Hospital organization culture, capacity to innovate and success in technology adoption. *J Health Organ Manag* 2006 May;20(2-3):194-217. [doi: [10.1108/14777260610662735](https://doi.org/10.1108/14777260610662735)] [Medline: [16869354](https://pubmed.ncbi.nlm.nih.gov/16869354/)]
26. Bardach N, Huang J, Brand R, Hsu J. Evolving health information technology and the timely availability of visit diagnoses from ambulatory visits: a natural experiment in an integrated delivery system. *BMC Med Inform Decis Mak* 2009 Jul 17;9:35 [[FREE Full text](#)] [doi: [10.1186/1472-6947-9-35](https://doi.org/10.1186/1472-6947-9-35)] [Medline: [19615081](https://pubmed.ncbi.nlm.nih.gov/19615081/)]
27. Murovec N, Prodan I. Absorptive capacity, its determinants, and influence on innovation output: Cross-cultural validation of the structural model. *Technovation* 2009 Dec;29(12):859-872. [doi: [10.1016/j.technovation.2009.05.010](https://doi.org/10.1016/j.technovation.2009.05.010)]
28. Kotlar J, De Massis A, Frattini F, Kammerlander N. Motivation gaps and implementation traps: the paradoxical and time - varying effects of family ownership on firm absorptive capacity. *J Prod Innov Manag* 2019 Aug 30;37(1):2-25. [doi: [10.1111/jpim.12503](https://doi.org/10.1111/jpim.12503)]
29. Cohen W, Levinthal DA. Absorptive capacity: a new perspective on learning and innovation. *Administrative Science Quarterly* 1990 Mar;35(1):128. [doi: [10.2307/2393553](https://doi.org/10.2307/2393553)]
30. van der Meij EVA, Bouwsma EVA, van den Heuvel B, Bonjer HJ, Anema JR, Huirne JAF. Using e-health in perioperative care: a survey study investigating shortcomings in current perioperative care and possible future solutions. *BMC Surg* 2017 May 23;17(1):61 [[FREE Full text](#)] [doi: [10.1186/s12893-017-0254-6](https://doi.org/10.1186/s12893-017-0254-6)] [Medline: [28535763](https://pubmed.ncbi.nlm.nih.gov/28535763/)]
31. Lenox M, King A. Prospects for developing absorptive capacity through internal information provision. *Strat. Mgmt. J* 2004 Apr;25(4):331-345. [doi: [10.1002/smj.379](https://doi.org/10.1002/smj.379)]
32. Greenhalgh T, Robert G, Macfarlane F, Bate P, Kyriakidou O. Diffusion of innovations in service organizations: systematic review and recommendations. *Milbank Q* 2004;82(4):581-629 [[FREE Full text](#)] [doi: [10.1111/j.0887-378X.2004.00325.x](https://doi.org/10.1111/j.0887-378X.2004.00325.x)] [Medline: [15595944](https://pubmed.ncbi.nlm.nih.gov/15595944/)]
33. Ahmadi H, Sadoughi F, Gholamhosseini L, Azadi T, Sheikhtaheri A, Nilashi M, et al. Organizational factors affecting the adoption of a technological innovation: a study of the hospital information system (HIS). 2019 Presented at: 2019 5th International Conference on Web Research (ICWR); April 24-25, 2019; Tehran, Iran. [doi: [10.1109/icwr.2019.8765270](https://doi.org/10.1109/icwr.2019.8765270)]
34. Jeyaraj A, Rottman JW, Lacity MC. A review of the predictors, linkages, and biases in IT innovation adoption research. *Journal of Information Technology* 2006 Mar 01;21(1):1-23. [doi: [10.1057/palgrave.jit.2000056](https://doi.org/10.1057/palgrave.jit.2000056)]
35. Rathert C, Williams E, Linhart H. Evidence for the quadruple aim: a systematic review of the literature on physician burnout and patient outcomes. *Med Care* 2018 Dec;56(12):976-984. [doi: [10.1097/MLR.0000000000000999](https://doi.org/10.1097/MLR.0000000000000999)] [Medline: [30339573](https://pubmed.ncbi.nlm.nih.gov/30339573/)]
36. van Andel J, Leijten F, van Delden H, van Thiel G. What makes a good home-based nocturnal seizure detector? A value sensitive design. *PLoS One* 2015;10(4):e0121446 [[FREE Full text](#)] [doi: [10.1371/journal.pone.0121446](https://doi.org/10.1371/journal.pone.0121446)] [Medline: [25875320](https://pubmed.ncbi.nlm.nih.gov/25875320/)]

Abbreviations

ACWO: Advisory board for Science and Research

CONSORT-EHEALTH: Consolidated Standards of Reporting Trials of Electronic and Mobile HEalth Applications and onLine TeleHealth

COREQ: consolidated criteria for reporting qualitative research

eHAM: eHealth adoption model

EMR: electronic medical record

IT: information technology

PDMS: patient data management system

Edited by A Kushniruk; submitted 20.09.21; peer-reviewed by AR Mohseni, B Chaudhry; comments to author 07.12.21; revised version received 13.01.22; accepted 03.02.22; published 26.04.22.

Please cite as:

Thiel B, Iao I, Smid J, de Wit E, Koopman S, Geerts B, Godfried M, Kalkman C

Adoption of a Postoperative Pain Self-Report Tool: Qualitative Study

JMIR Hum Factors 2022;9(2):e33706

URL: <https://humanfactors.jmir.org/2022/2/e33706>

doi: [10.2196/33706](https://doi.org/10.2196/33706)

PMID: [35471472](https://pubmed.ncbi.nlm.nih.gov/35471472/)

©Bram Thiel, Inez Iao, Joris Smid, Emmy de Wit, Seppe Koopman, Bart Geerts, Marc Godfried, Cor Kalkman. Originally published in JMIR Human Factors (<https://humanfactors.jmir.org>), 26.04.2022. This is an open-access article distributed under the terms of the Creative Commons Attribution License (<https://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Human Factors, is properly cited. The complete bibliographic information, a link to the original publication on <https://humanfactors.jmir.org>, as well as this copyright and license information must be included.

Viewpoint

Toward an Ecologically Valid Conceptual Framework for the Use of Artificial Intelligence in Clinical Settings: Need for Systems Thinking, Accountability, Decision-making, Trust, and Patient Safety Considerations in Safeguarding the Technology and Clinicians

Avishek Choudhury¹, PhD

Industrial and Management Systems Engineering, Benjamin M Statler College of Engineering and Mineral Resources, West Virginia University, Morgantown, WV, United States

Corresponding Author:

Avishek Choudhury, PhD

Industrial and Management Systems Engineering
Benjamin M Statler College of Engineering and Mineral Resources
West Virginia University
1306 Evansdale Drive
PO Box 6107
Morgantown, WV, 26506-6107
United States
Phone: 1 5156080777
Email: avishek.choudhury@mail.wvu.edu

Abstract

The health care management and the medical practitioner literature lack a descriptive conceptual framework for understanding the dynamic and complex interactions between clinicians and artificial intelligence (AI) systems. As most of the existing literature has been investigating AI's performance and effectiveness from a statistical (analytical) standpoint, there is a lack of studies ensuring AI's ecological validity. In this study, we derived a framework that focuses explicitly on the interaction between AI and clinicians. The proposed framework builds upon well-established human factors models such as the technology acceptance model and expectancy theory. The framework can be used to perform quantitative and qualitative analyses (mixed methods) to capture how clinician-AI interactions may vary based on human factors such as expectancy, workload, trust, cognitive variables related to absorptive capacity and bounded rationality, and concerns for patient safety. If leveraged, the proposed framework can help to identify factors influencing clinicians' intention to use AI and, consequently, improve AI acceptance and address the lack of AI accountability while safeguarding the patients, clinicians, and AI technology. Overall, this paper discusses the concepts, propositions, and assumptions of the multidisciplinary decision-making literature, constituting a sociocognitive approach that extends the theories of distributed cognition and, thus, will account for the ecological validity of AI.

(*JMIR Hum Factors* 2022;9(2):e35421) doi:[10.2196/35421](https://doi.org/10.2196/35421)

KEYWORDS

health care; artificial intelligence; ecological validity; trust in AI; clinical workload; patient safety; AI accountability; reliability

Introduction

With the growth of multidisciplinary and collaborative health care [1], clinicians have more information and expertise to inform clinical decision-making than ever before [2]. Nevertheless, when confronted with information and knowledge that are (1) not always within the scope of the primary or focal expertise of a clinician and (2) in such quantities that it becomes

difficult for the clinician to process reliably and validly and in a timely manner, clinicians can often resort to boundedly rational and, in some cases, incorrect diagnoses, treatment, and other clinical decisions [3]. A response to the interrelated problems of the clinician's limited absorptive and cognitive capacities has been the integration of artificial intelligence (AI) into health care decision-making [4-6]. However, technological solutions to the problem of limited absorptive and cognitive capacities in multidisciplinary, complex, and collaborative decision-making

can introduce new situations [7]. For example, *team science* in clinical settings can come with competing diagnoses and prescriptions for treatment and wellness [8,9]. Furthermore, when new technologies for decision-making are imposed *from above* (eg, by management, rather than organically) or *from below* (eg, at the clinician level, clinicians may not always trust or intend to use those technologies) [10].

The problems regarding trust in AI and the use of AI systems in clinical decision-making illustrate the classic distinction between the rational and descriptive decision-making models. Studies of clinical decision-making demonstrate that the rational model of introducing integrative technologies, including but not limited to AI, into clinical decision-making is not always supported by the data. In other words, rational models of clinical decision-making [6,11,12] and decision-making, in general, are not ecologically valid; they assume perfect information, ideal absorptive and cognitive capacity, optimal trust, and unlimited resources to make a fully and correctly informed decision. The descriptive empirical research demonstrates mixed effects regarding technology-assisted decision-making in clinical settings owing to limited cognitive capacity of the end user (care providers), information overload or lack of data, and suboptimal trust in the technology [7,13,14].

Similar to most technologies, AI can be a boon or bane within the health care ecosystem. With increasing autonomous activities in health care, challenges concerning AI and human factors may manifest evidently at an individual level (eg, awareness and trust), macrolevel (eg, regulation and policies), and technical level (eg, usability and reliability) because many health care AI applications are poorly designed and not evaluated thoroughly [15]. Therefore, human factors and ergonomics (HFE) consideration in health care AI systems has become necessary. If leveraged while developing AI systems, HFE principles and methods can augment its use and adoption without disturbing patient safety or clinical protocols. Of all the possible HFE challenges that AI in health care can cause, suboptimal clinician-AI interaction is significant. Integration of poorly designed AI in health care can complicate the relationships

between clinicians and computer (intelligent) systems. Unlike other health care technologies, the complexity of AI is more, as it can interact (through chatbots, automated recommender systems, health apps, etc) with clinicians and patients based on the inputs (feedback) that it receives from them. AI's output (result generated by the AI) largely depends on the information fed into it—certain types of AI, for instance, reinforcement learning [16], learn and adapt themselves based on user input to optimize the outcome. Therefore, clinician-AI interaction may influence AI performance and, in turn, the clinician's viewpoint toward it. Optimal and successful clinician-AI interaction depends on several factors, including situation awareness, cognitive workload, working environment, and emotional resources (eg, current state of mind, willingness to use AI, previous experience with AI technology, trust in technology, and others). Most studies on health care AI have ignored (1) ecological validity and (2) human cognition, which may create challenges at the interface with clinicians and the clinical environment. Moreover, there is a lack of sufficient studies focusing on improving the human factors, mainly, (1) how to ensure whether clinicians are implementing it correctly; (2) the cognitive workload it imposes on clinicians working in stressful environments; and (3) its impact on clinicians' situation awareness, clinical decision-making, and patient safety outcome. Although studies on AI have reported its great performance and potential in medicine [17-19], research breakthroughs (AI performance in research settings) do not necessarily translate into a technology that is ready to be used in a high-risk environment [20], such as health care. In addition, most AI featuring prominent abilities in research and literature are not executable in a clinical environment [21,22]. According to the technology readiness level (TRL), most AI systems, at least in pediatric and neonatal intensive critical care, if not all, do not qualify for implementation [17].

TRL is a gauging system, developed to assess the maturity level of a particular technology [23]. It consists of 9 categories (readiness levels), in which a score of TRL 1 is the lowest and TRL 9 is the highest (Textbox 1).

Textbox 1. Technology readiness levels (TRLs; 1-9).

Technologies with TRL 1-4 are executable in a laboratory setting, where the main objective is to conduct research. This stage is the proof of concept.

- TRL 1: Basic principles of the technology observed
- TRL 2: Technology concept formulated
- TRL 3: Experimental proof of concept developed
- TRL 4: Technology validated in a study laboratory

Technologies with TRL 5-7 are in the development phase, in which the functional prototype is ready.

- TRL 5: Technology validated in a relevant environment (controlled setting in a real-life environment)
- TRL 6: Technology demonstrated in a relevant environment
- TRL 7: System prototype demonstrated in an operational environment

Finally, technologies with TRL 8 and 9 are in the operational phase, in which the primary objective is implementation.

- TRL 8: System completed and certified for commercial use
- TRL 9: System approved for and implemented in the actual operational environment

For example, in clinical settings, nurses and physicians have demonstrated lack of trust in AI, including machine learning analytics and decision-making tools [7]; numerous other information technologies designed to improve decision-making efficiency and effectiveness, such as medication management systems [13], event reporting systems [14], and electronic health care records systems [24]; and clinical biotechnologies such as gene therapy [25]. There are demonstrations of incorrect use of clinical technologies, such as unwarranted trust and reliance on automated nursing tools, leading to adverse health consequences, including, but not limited to, avoidable fatalities [11] and inappropriate use of medical devices inducing patient harm [12,26]. It is critical to understand that the impact of AI, particularly in health care, is not only a function of the accuracy of its underlying mathematical process but also the cognitive human factors, including trust, perception, usability, and safety. Therefore, to minimize errors caused by health care AI (as noted in other health information technology [HIT] literature, such as electronic medical records), a holistic approach, recognizing health care as a dynamic sociotechnical system in which subelements interact, is necessary.

Objective

This study aimed to propose a descriptive conceptual framework derived from cognitive human factors and decision-making literature. Note that this framework is not a rational model. Then, future studies can leverage this framework and inform the eventual development of a prescriptive framework for optimal AI-clinician interactions. The proposed framework can be best used for mixed methods studies. In other words, the descriptive conceptual framework will help to capture the interactions between clinicians and AI. The prescriptive framework (guided by experimental study findings) will help to develop better AI-clinician interactions.

The novelty of the descriptive framework presented in this study is that it uses systems thinking and combines multiple descriptive (vs rational) human factors approaches to understand clinician-AI interactions in decision-making. Although human factors considerations in clinical decision-making can augment the intended positive impacts of integrative decision-making technologies such as AI, so far, there are few studies on how and the extent to which clinicians use AI in diagnostic and health care decision-making. In addition, the predominance of empirical studies of AI in clinical settings focuses on the technical aspects of AI-driven diagnostic and care decision-making, that is, the plethora of machine learning algorithms and high-dimensional data that AI entails [27]. The few studies on human factors in the use of AI in decision-making are not focused on clinical samples and contexts, but rather on nonclinical applications in other industries and sectors of the economy, including, but not limited to, military [28]; transportation [29]; and organizational design, in general [30].

Framework Development

The health care management and medical practitioner literature lack a conceptual framework for capturing the impact of AI from a systems perspective and simultaneously understanding

clinician-AI interactions that are ecologically valid, specifically focusing on how such interactions may vary based on human factors such as expectancy, trust, cognitive variables related to absorptive capacity and bounded rationality, and concerns for patient safety. To derive the conceptual framework, this study leverages (1) literature on systems thinking and AI in medical practice, (2) information use in human decision-making, (3) trust and informing decisions with AI, and (4) patient safety and informing decisions with AI.

Systems Thinking and AI in Medical Practice

Overview

Technological advancement and diffusion of innovation are supporting an expeditious transformation in the structures and institutions in veritably every facet of life, and medical practice is no exception. Technologies can now facilitate the accomplishment of activities that humans once considered impossible and are responsible for substantial social and public policy changes in health care. For example, Widmer et al [31] discussed the convergence of health care policy reform in the United States with technological advancements and social shifts as support for the great use of AI in health care practice. They argued that these are transformational forces that influence the capacity to develop complex solutions to problems in medicine. These solutions are in the form of technologies that often rely on AI to support decision-making. Qadri et al [32] surveyed the current landscape of new health care technologies, uncovering the ubiquity of AI and tools dependent on AI in medicine. For example, the impact of the health care Internet of Things on health care information technology has been substantial [32], as the immensity of technological innovation relentlessly pushes forward as systems become increasingly smart and widespread. As these systems become an integral part of health care, systems thinking will become increasingly essential because of the complex nature of the task-technology fit required in health care.

The health care industry has witnessed several design errors in both technologies and clinical workflow. Integrating HITs that are not designed and not tested properly is highly likely to contribute to new categories of technology-induced errors, often new to the health care domain. Such errors usually manifest in the complex interaction between health care providers and HIT during actual clinical use. For example, in the recent past, surgical robots were responsible for 144 patient deaths and 1391 patient injuries [33]. Once integrated, such technologies can also alter the existing clinical workflow. For example, integration of AI into the clinical workflow without considering its impact on clinicians, patients, hospital expenses, workflow speed, insurance claiming process (previous authorization), and other aspects can disrupt the overall care process. For example, given the dependence of AI on data, it is feasible to assume that even the best AI systems will sometimes be wrong, leading to compromised patient safety. Although clinical errors and near misses are common in health care, AI errors can be significantly unique. First, errors arising from AI systems can become widespread without being identified by clinicians, causing system-wide error—rather than the limited number of patients injured by any provider's error. Second, tracking AI errors can

also become highly challenging, mainly when powered by deep learning algorithms. Such a complex system (AI) can make root cause analysis very daunting and almost impossible owing to its inherent opaque nature. The performance of AI systems largely depends on the data on which they have been trained. As the existing data repositories are biased, AI integration without addressing issues regarding data quality can escalate health care biases.

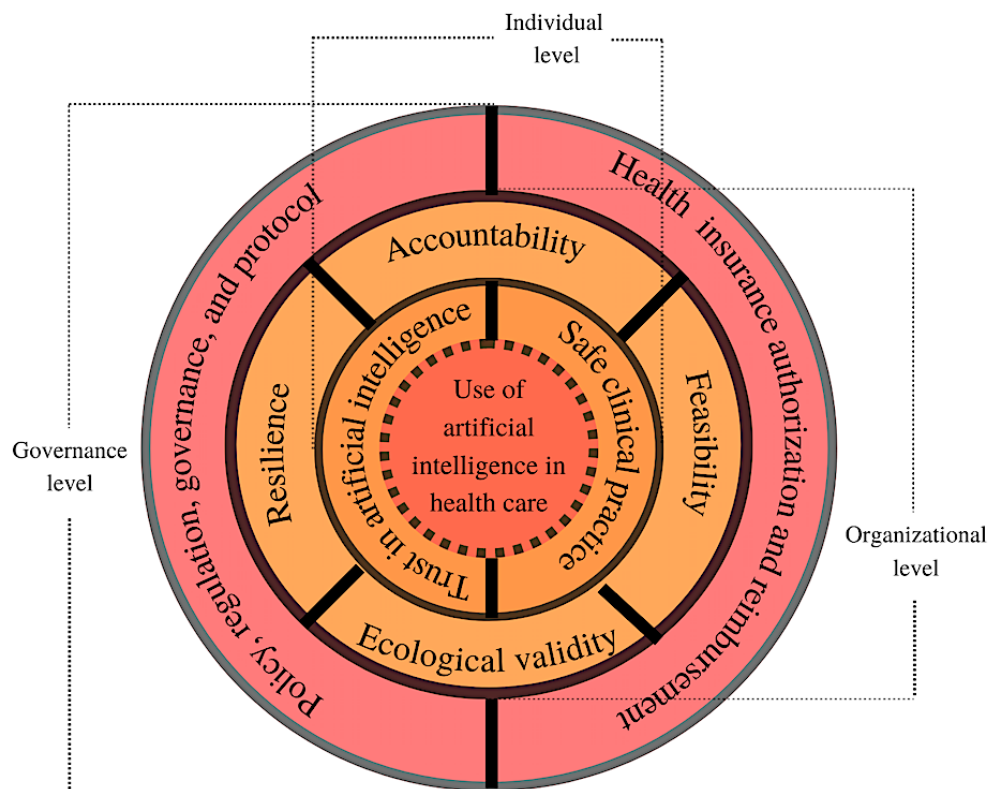
Health care authorities must account for several extrinsic factors such as clinicians' willingness to use AI in their clinical practice, access to the duration and frequency of AI training required by clinicians with different expertise, and feasibility of personalizing AI for individual clinicians and patients. In addition, doctors and nurses, the potential users of AI in a hospital, can also misuse the system either owing to lack of AI

literacy or poor usability of AI. Therefore, a systems thinking approach is essential for the safe integration of AI.

In addition, AI-based technologies may not work well for patients with rare diseases, as their data are not adequately available. Health care authorities will also have to ensure that, over time, clinical experts do not become deskilled or permanently replaced owing to AI implementation. In other words, safe and sustainable integration of AI requires a systems approach in which all interactions between different health care stakeholders are considered.

Similar to any complex system, subsystems of health care and AI can be shaped by several factors at three major levels: (1) governance—policies, regulations, and protocols; (2) organizational [34]—accountabilities, resilience, ecological validity, and feasibility; and (3) individual—trust in AI and safe practices (Figure 1).

Figure 1. Factors influencing the use of artificial intelligence in health care—a systems viewpoint.



Governance Level

In this study, AI governance has been defined as a group of systems that regulate and control AI within the large health care ecosystem. It steers organizational objectives and risk monitoring to achieve optimized performance. In other words, AI governance is a *system of systems* that requires a holistic approach, incorporating strategic planning at all organizational levels. Existing studies have confined health care AI governance within the boundaries of organizational structure and processes for clinical decision-making, transparency without exploiting proprietary rights, fairness of the technology, and accountability [35].

Nevertheless, many critical factors have not been considered. Resilience; ecological validity; protocols for safe practices using AI; engagement; and responsibilities of stakeholders, including insurance providers; and human factors should also be included as significant components of health care AI governance. *Systems thinking* in health care can help regulatory authorities and organizations to perceive the integration of AI and health care as a merger between 2 complex systems. In other words, a systems approach will allow us to capture and understand how the dynamic relationships between various factors, such as policies and protocols, impact the resilience and feasibility of the incorporation of AI into the health care ecosystem. Clearly defined policies and protocols and involvement of all

stakeholders will also help to resolve the current concerns regarding AI accountability—who should be responsible for a flawed AI system or incorrect AI output? From a human factors perspective, the systems approach can advocate for the ecological validity of AI, ensuring that the technology is appropriately designed and developed for a given uncontrolled environment. Addressing these concerns can increase the likelihood of AI acceptance among clinicians, by augmenting their initial trust in the technology.

Organizational Level

Further expanding on the subcomponents of [Figure 1](#), AI governance in health care should account for (1) resilience thinking approach, (2) accountability, and (3) ecological validity of AI. *Resilience thinking* is a holistic way of investigating how interacting systems of clinicians and clinical environment, including AI technologies, can be best managed during uncertainty or systemic errors.

This study defines *AI accountability* as a process in which health care practitioners have potential responsibilities to justify their

clinical actions to patients (or families) and are held liable for any impending positive or negative impact on patient health. While using an AI-based decision support system, only clinicians are held accountable if they decide to follow AI, resulting in patient harm. Clinicians are also held responsible if they deviate from the standard protocols [36]. This may be worrisome because, under such circumstances, clinicians will only follow AI if it matches their judgment and aligns with the standard protocol—making the AI underused. According to our recent survey (*institutional review board 2022-007* approved by the Stevens Institute of Technology, the United States) consisting of 265 clinicians actively practicing in the United States, lack of AI accountability is a significant hindrance to AI adoption in health care. Clinicians hesitate and refuse to use AI as they do not want to take responsibility for faulty AI. Participants in our survey advocated for contractual agreements with patients and policies to safeguard them from AI errors and related patient safety issues. [Textbox 2](#) shows some of the responses provided by clinicians.

Textbox 2. Solutions provided by health care practitioners to address the lack of artificial intelligence (AI) accountability—categories and sample responses.

Contractual agreement with patients

- “I think a solution would need to be having patients sign informed consent for AI to be used in their care and that decisions made by the AI cannot reflect on the provider’s care.”
- “AI should only be used if a patient is willing to fill out a questionnaire regarding the pros and cons of using AI and the potential harm or good, releasing the practitioner along with the potential outcome they may or may not achieve.”
- “[I] think the patient should sign a waiver if AI is used.”

Policies and safety measures

- “Use it in conjunction with training and safeguards that are in place now.”
- “The use of AI would need to be regulated. The manufacturers should take full responsibility for any negligent or bad decisions about patient care.”
- “[I] would not want to be held accountable for AI recommendations. Creating policies to protect clinicians would be important.”

Being a complex system, subsystems of health care establishments are shaped by several internal and external factors. This complexity of the health care system can be well explained by using human factors approaches such as the Safety Engineering Initiative for Patient Safety (SEIPS) framework [37]. Developed by Carayon et al [37], the SEIPS framework is partly based on the well-known structure-process-outcome model of health care quality by Donabedian [38]. It is arguably one of the most acknowledged and published systems-based human factors frameworks in health care. SEIPS framework illustrates the dynamic interactions between people (patients and clinicians), technology (AI in this context), tasks (clinical activities to support patient safety and health), and environment (clinical and organizational setting). However, no studies have used the SEIPS framework to understand the impact of AI on health care from a systems perspective.

Individual Level

Diligent scrutiny is essential for medical practitioners when considering the application of new technologies in patient care. There are limitations to the benefits of AI in medical practice.

Failing to acknowledge them when engaging in innovative decision-making, especially when human lives are at risk, can result in system accidents. Research [39] has discussed the limitations attached to AI application in medicine, focusing on its application in oncology. They noted that machine learning plays a substantial role in oncological practice. Machine learning, which is a subset of AI, involves computers’ ability to learn autonomously through data input [39]. In oncology, benefits of machine learning include application in risk modeling, engaging in diagnostic and staging investigation, prognosis prediction, and therapy response prediction. Limitations persist when using AI, such as costs, overdependence on data quality, black box effect, and obtaining trust in and acceptance of machine learning technology [39]. Mendelson [40] echoed some of the limitations discussed by Khan et al [39], noting that physicians cannot rely on AI alone when making decisions about the findings from breast imaging examinations. Mendelson [40] described the preferred role of AI as being supportive of diagnosis and patient management.

Although physicians and researchers describe several limitations to accepting AI owing to its analytical abilities and biases, other human factors have often been neglected. Although the refinement of methods and procedures used in AI for decision-making continues to advance, further exploration of leveraging human factors principles is mandatory. The solution to safeguard AI and patients is in the acceptance of systems thinking approach to medical care, in which physicians incorporate AI in a role that is, as Mendelson [40] noted, supportive in nature. As described by Khan et al [39], the problem of trust with AI appears to be well founded because of the black box effect, in which the AI delivers results with a solution; however, the rationale for the solution cannot be described. Hashimoto et al [41] noted that although the black box effect exists, efforts are underway to design solutions that can mitigate it in medical practice. The black box results obtained from neural network methods can lead to the correct response [41]. However, physicians cannot rely on the results from AI alone at this point, when making decisions that have life-or-death implications for patients. The system cannot explain itself. Although physicians can learn more about AI to understand the results better [40], the problem of human trust in AI remains challenging because the design of AI does not support transparency [40]. Hence, *the limitations of AI are not based on AI alone, but they are based on the relationship between AI and its users' lack of understanding of the technology. Not all the trust needed to rely on AI can come from better design features. Some must come from great acceptance of the technology, and the interdisciplinary nature of systems thinking can play a role in improving the relationship between humans and AI in medicine.* However, it is essential to keep in mind the dynamic nature of trust, where the user needs to have a priori trust in technology to use it for the very first time. Moving forward, their trust in it can become a function of their experience with the technology and its effectiveness.

In addition to the possibility of patient harm caused by disruption to health care delivery, the complexity of how systems fit together can result in system accidents. Kappagoda [42] discussed the problem of system accidents in aviation to illustrate the potential for problems when there are design deficiencies, poor maintenance practices, and failures in oversight. Similarly, poor AI design in health care can lead to patient harm, where clinicians can misinterpret AI information or click on the wrong option on the AI display. Inadequate maintenance, that is, not retraining the AI with new patient data, can compromise its prediction accuracy, and thus hinder patient safety [43]. In addition, sometimes, bedside care providers make clinical decisions that do not necessarily fall within the standard guideline (for specific patient types) or skip the prescribed clinical steps (under excessive workload) to accomplish a particular clinical goal promptly [44]. Therefore, AI developers should account for such human behavior while designing their products, so that AI can serve as a support and not as a hurdle in the everyday clinical workflow.

Some assurances related to medical devices exist in the United States. These include the International Organization for Standardization 13485 quality standards for medical devices [45] and 21 Code of Federal Regulations 820.3(l) [46]. Although

these certifications and regulations exist to protect patients, medical personnel can still preside over a case in which system accidents harm a patient. Hence, although AI can be substantially beneficial to patients and be a helpful tool for supporting the staff's decisions, medical professionals must engage in systems thinking when assessing care strategies.

Information Use in Human Decision-making

The motivation to use or not use certain information in decision-making is complex, and several theoretical perspectives, such as Situation Awareness and Expectancy Theory, can support the understanding of this motivation. Soltani and Farhadpour [47] investigated user motivation toward using information services. The framework for their study was the expectancy theory. They found that user motivation to use an information service was significantly influenced by awareness of results value and perceptions of the accessibility [47]. *Although expectancy theory appears to play a role in describing why AI is used or not used, other human factors can support in predicting user behavior.* O'Reilly [48] examined the variations in the use of information sources to understand the impact of quality and accessibility of information as factors influencing its use. O'Reilly [48] found that the frequency of use was the most significant influencer of use.

The association of absorptive capacity with the ability to use information is another essential facet of the psychology of decision-making. Results from the study by Liao et al [49] indicate that absorptive capacity has an impact on innovation, but information use aimed at innovation was found to be complicated [50]. Schmidt [50] discovered that the determinants of absorptive capacity are different, depending on the type of knowledge absorbed. Therefore, absorptive capacity is path-dependent in how it leads to information use. This is a complexity that constrains decision-making. In decision-making research, perceived relevance of and access to information are critical to understanding information use. One of the first studies to understand the effects of information relevance on decision-making in complex environments was by Streufert [51]. The framework for her study was the complexity theory. She noted information relevance as a factor that affected complex decision responses, but the same element (information relevance) failed to influence simple decision responses [51]. These findings are critical to understanding the significance of information in decision-making research because they support the essential nature of *situation awareness* among decision makers. *Her conclusion that complex decision-making is affected by relevance and simple decision-making is affected by information load—is a critical finding,* placing some limitations on the complexity theory. Citroen [52] explored the role of information in strategic decision-making by executives in organizations. The approach requires executives to collect and use information in a structured process that supports the elimination of uncertainty in the decision-making process. The findings of Streufert [51] and Citroen [52] support the role of situation awareness as a factor that influences information use among decision makers.

The acceptance of information in decision-making is another pivotal factor in decision-making research. According to a

well-established model called technology acceptance model, acceptance is associated with ease of use and usefulness in decision-making [53]. The inclusion of information in decision-making would appear to be important. However, its inclusion and tools such as decision support systems remain as a challenge for decision makers. Todd and Benbasat [54] examined the use of information in decision-making. Their study was critical of assuming that managers who have more information will make better decisions. They found that *the conservation of effort* occurred when managers were presented with more information. Tools such as decision support systems did not result in a great likelihood of information being used in the decision-making [54]. These findings can have substantial implications for studies on human decisions formed by AI, because AI often involves aggregating several piles of data to construct a comprehensive understanding of the phenomenon under investigation. However, if what Todd and Benbasat [54] proposed is consistent with decision makers' current approach to information, the aggregation of data to create elegant models to understand a phenomenon will go in vain. Studies exist on why individuals choose to rely on information systems for decision-making. Snead and Harrell [55] examined the decisions by the management to use decision support systems using expectancy theory. The findings indicate that the expectancy force model can determine managerial behavioral intentions to use decision support systems [55]. Behavioral theory is helpful for these studies because it can support the assessment of why people use systems without previous experience with them based on intention and motivation.

Absorptive capacity is associated with the use of AI, and the findings from AI are critical. Absorptive capacity is also essential in decision-making related to innovation and depends on how a user optimizes information system capabilities. Moreover, absorptive capacity also impacts AI in terms of industry innovation. *A limitation of AI use is the lack of user understanding of tools or substantively interpreting findings.* Shi et al [56] discovered that AI use creates challenges in terms of limitations, such as limited knowledge transfer. The extent to which workers are trained to use AI tools and interpret their findings is limited. Therefore, absorptive capacity in business settings where AI is used is limited by workers' capabilities.

The knowledge and relevance of AI are also essential to consider while supporting decision-making. Prevedello et al [57] examined the challenges of AI use in medical settings. They noted a difference between expectations and AI application in clinical settings where AI's role in tasks, such as radiology, would expeditiously advance purely from a technical standpoint without addressing all the user needs from a human factors perspective. Prevedello et al [57] noted that AI should be a part of developing clinically relevant outcomes and that AI should play a role in decision-making in the future. However, this is also a prediction that Prevedello et al [57] found to have gone unfulfilled from previous studies. Pomerol [58] discussed the issue of AI and human decision-making. He described AI as sharing several relationships with other types of quantitative analytical procedures in that each is useful in diagnosis. He also noted that a critical limitation of AI was the lack of capacity

for look-ahead reasoning, where uncertainty and preferences are crucial factors to consider [58].

Acceptance of AI in decision-making is a critical technological concept in which the ease of use and usefulness of AI is examined and determined. The use and benefit of AI in decision-making are substantially challenged by lack of knowledge of the technology or its potential capabilities. Chan and Zary [59] discussed the applications and challenges of AI implementation in medical education. One of the major factors restraining AI use in the medical profession is that the medical school curriculum fails to develop future medical professionals to understand AI algorithms [59]. The lack of knowledge and development results in limited use of the tool. A critical limitation to the use of AI in the future appears to not be caused by the constraints or complexities of the technology, but instead by the decision to use the technology by humans [59]. Sohn and Kwon [60] examined several technology acceptance theories to understand which framework best fits the acceptance of AI. Their study included the technology acceptance model, theory of planned behavior, unified theory of acceptance and use of technology, and value-based adoption model. The findings supported the value-based adoption model as the best model to determine user acceptance of AI. Specifically, the factors found to have the most significant impact were *enjoyment* and *subjective norms* [60]. These findings provide evidence that the motivation to use AI is driven more by interest in technology than the utilitarian aspects of AI.

Trust and Informing Decisions With AI

Trust in technology is influential in several contexts, including those where computer-mediated communication is used for work team communication [61], supporting customers or clients engaging in electronic channels, e-commerce [62], and aviation activities [63]. Trust in technology delineates from trust in humans in many different ways [64,65], in that trust in humans is associated with interpersonal relationship qualities. In contrast, trust in technology is associated with reliability and performance. Nevertheless, trust remains an important aspect of the human experience in technology [65].

Trust in technology appears to be consistent with theories such as the expectation disconfirmation theory. This theory is related to the satisfaction an individual has with experience related to whether their beliefs were confirmed during an experience and how expectations and perceived performance affected their initial beliefs [66]. Trust in technology is complex for many reasons. A reason for the complexity of trust in technology is that there are risks and uncertainties associated with technology. Li et al [67] examined trust in new technology in the context of the workplace. They found that initial trust formation relies on several factors, including trusting bases, beliefs, attitudes, organization's subjective norms, and trusting intentions. Other studies involving technology assume that trust in technology can be formed through governance in the organization. Winfield and Jirotko [34] discussed the development of a framework for ethical governance, pertaining to robotics and AI systems in organizations. Factors that were considered included ethics, standards, regulation, responsible research, innovation, and public engagement. These factors were deemed essential in the

development of trust between the technology and public. The problem with this approach is that it does not consider human factors such as users' perception of technology, perception of risk associated with it, and its impact on users' cognitive workload and situation awareness.

Most relevant to this research is the issue of trust in medical technology. Montague et al [68] examined trust in medical technology and sought to describe medical technology as a distinct construct from trust in general technology. A review of the literature on trust in technology was included in their study. Their literature discussion included the assertion that previous study findings support a lack of difference between trust in humans and trust in technology. However, McKnight et al [64] and Lankton et al [65] have included findings and discussion, which indicate a substantial difference between trust in humans and trust in technology. The difference in the findings supports further investigation. The existence of an entirely separate construct describing or measuring trust in medical technology, aside from trust in general technology, should not be considered in subsequent studies to conclude whether a separate construct exists. However, it is beyond the scope of this study.

Specifically, trust in AI remains as an important issue and will grow significantly with time as AI becomes increasingly infused into the products we use in everyday life. AI continues to create some difficulty among researchers regarding how AI should be trusted. AI uses large amounts of data to support decisions that receive attention and consideration based on strong predictability, while not mimicking humans' thought processes. Hurlburt [69] discussed AI as a technology that continues to increase its reach and that people have become increasingly dependent on the use of AI in their everyday lives. The problem is that, often, there is lack of consideration as to whether AI is capable of doing the job it was selected to perform.

Furthermore, vulnerabilities with AI continue to persist. Hurlburt [69] noted that AI should be trusted only to a certain extent. However, the consideration and act of trust toward selecting AI to complete tasks is better suited for the individual evaluating the AI tool than the task itself.

We also advocate for some level of skepticism regarding the decisions made by AI. The amount of skepticism necessary for the most accurate clinical decisions depends on the capability of the clinician and AI system. Suppose the benefits and constraints of an AI tool are understood. In that case, decision-making about whether to use the tool entails placing trust (binary in nature), rather than considering the extent to which a tool with predefined specifications can be trusted. The authors' effort to analogize trust in AI with trust in humans is remarkable in the literature on AI involving trust. For example, in the study by Hengstler et al [70], trust in AI used in tools such as autonomous vehicles and medical assistance devices was investigated [70]. They sought to draw an analogy between applied AI in vehicles and human social interaction [70]. Their focus was on understanding the relationship between humans and automation, to understand how trust is built. They concluded that trust in AI is inextricably linked to the trust that individuals have in the firm that created the AI. The philosophical approach to this research is very different from that of researchers

examining trust in technology in general [64,65]. The focus of researchers was on establishing that trust in technology and trust in humans are entirely different concepts. Therefore, the nature of AI as a form of intelligence designed to be similar to human intelligence can affect how AI is considered, even in scholarly research.

People's trust in AI shares commonalities between trust formation in automation (non-AI technology) and interpersonal trust (trust in humans). Glikson and Woolley [71] discussed previous studies involving human trust in AI. They noted that there are critical differences between AI and other technologies, which impact how trust forms and works, similar to both human and technology trust. Cognitive and emotional trust in AI are related to both the representation of the AI, whether robotic, web-based, or embedded, and the level of intelligence of the AI system. These factors are integral to establishing people's cognitive and emotional trust in AI. In the scope of AI use in health care, the conceptualization of AI as having some anthropomorphic qualities becomes increasingly visible. Kerasidou [72] examined the use of AI in health care, focusing on the issues of empathy, compassion, and trust. She noted that these are characteristics that people should not value in AI. However, AI in medical treatment plays a role where AI completes several tasks that humans traditionally complete. The technology must be task-oriented and support humans in health care by performing more activities related to the emotional and comfort aspects of patients' treatment. Together, these findings contribute further support for AI, where the tool can fill a supportive role and enable humans to hold a position where trust would be beneficial to their health care delivery.

Patient Safety and Informing Decisions With AI

The most fundamental aspect of medical care is the promise of physicians to not harm (patient safety). The Hippocratic Oath is the standard that health care professionals must follow when working with patients. Therefore, understanding how AI impacts patient safety is critical for this study.

Health care AI studies have positively contributed to drug development, personalized medicine, and patient care monitoring [73-76]. AI has also been incorporated into electronic health records to identify, assess, and mitigate threats to patient safety [77]. Recent studies and reviews have primarily focused on the performance of AI at the diagnostic level, such as disease identification [78-83], and AI robotics in surgery and disease management [84-87]. Other studies have also implemented AI technologies to assist at the clinical level, including for assessing fall risks [88] and medication errors [89,90]. However, many, if not all, of these studies have not implemented AI in a clinical setting or have been used by clinicians for routine clinical activities. Therefore, we noted a lack of evidence that can confirm the positive impact of AI on patient safety outcomes in real life.

The impact of AI on patient safety substantially depends on how clinicians correctly comprehend AI output (information and recommendation) and accordingly make clinical decisions. In other words, misinterpretation of the AI output may mislead clinicians and encourage them to make wrong clinical decisions—putting patient safety at risk. With the integration

of AI, the role of technology shifts from merely *delivering information* to *information identification* and *decision-making*, therefore, enunciating the importance of clinician-AI interaction and collaborative decision-making. Most decision-making literature in the context of health care focuses on shared decision-making (clinician-patient) and its impact on patient safety. However, no studies have considered the significant role that AI can play in clinical decision-making (clinician-AI) and patient safety. Woolf et al [91] believed that an informed choice should occur in an interpersonal manner. Légaré et al [92] also discussed the importance of increasing the use of shared decision-making.

Nevertheless, the critical findings of health care decision-making literature may also apply to AI-based decisions. For instance, Edwards et al [93] found that in shared decision-making between the patient and clinician, the patient’s degree of health literacy determined their ability to understand their treatment (creating a shared mental model between the clinician and patient). Similarly, clinicians’ AI literacy will assess their ability to comprehend AI outcomes and make informed clinical decisions, thus ensuring treatment adherence and safety.

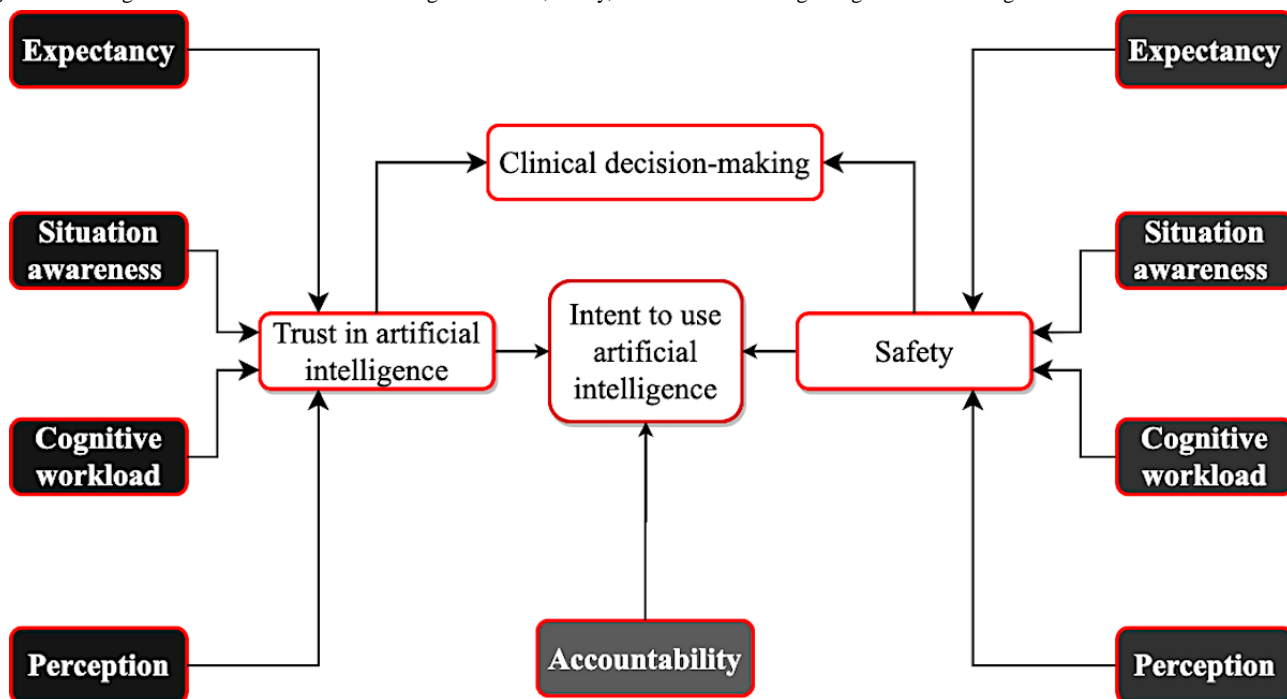
The importance of information interpretation and analysis, in general, has been well acknowledged in the literature [94]. For example, Tuffaha et al [95] discussed using the value of information analysis in health care as a model to support health care decision-making approaches. The value of the information

analysis approach supports the measurement of decision uncertainty and assessment of the evidence’s sufficiency to support technological implementation. Bindels et al [94] supported the use of the value of information analysis in health care decision-making. Although the value of information analysis is a practical approach to decision-making, there is a lack of implementation of AI and studies analyzing its impact on clinical decision-making and patient safety. These findings provide evidence that the issue of safety must receive further focus in the form of empirical research to inform patient safety and informed decisions regarding AI. The current body of research includes a rich collection of studies focused on using AI in tasks and decision support roles where the potential exists that users or those dependent on AI use are at risk of possible harm from AI technology.

The Proposed Framework

On the basis of the literature discussed previously, this study proposed the following conceptual framework (Figure 2). The framework emphasizes clinicians’ cognitive functions and perceptions regarding AI, concerning their trust in the technology along with perceptions of patient safety (risk). In addition, the framework emphasizes the cognitive functions of *situation awareness*, *workload*, *expectancy* (*performance and effort*), *trust*, *patient safety*, *clinicians’ perceptions of AI*, and *perception of AI accountability*.

Figure 2. Ecological validation of artificial intelligence—trust, safety, and decision-making using artificial intelligence.



To explore this conceptual framework for describing clinician-AI interactions in clinical decision-making, each independent variable has operational precedent in the cognitive human factors and behavioral economics literature. There are numerous measures of situation awareness, including, but not limited to, the 3-level model of Endsley, perceptual cycle model [96], and activity theory model [97]. Similarly, *workload* has numerous and moderately competing operationalizations based

on profession or occupation, including, but not limited to, scientific and clinical jobs and occupations [98,99]. Regarding *perceptions of AI*, there is relatively less precedent operationally [100,101]. Similar to the independent variables in the descriptive model of clinician-AI interactions, the dependent variables of *trust in AI* and *perceptions of patient safety* have numerous operationalizations across the computer science and health care literature [102,103].

Accordingly, the framework constitutes a sociocognitive approach that extends the theories of distributed cognition and, thus, accounts for the ecological validity of AI. The model leverages the measures that the studies reviewed in the previous sections imply (and often explicitly state), which must be included to understand the ecological validity of any model of human-AI interactions in decision-making. These validated and well-established scales include the modified National Aeronautics and Space Administration' task load index [104], extended unified theory of acceptance and use of technology model [105], multi-item and previously validated scales for trust [106], and Mission Awareness Rating Scale [107] for situation awareness [108]. Inherently, cognitive workload and situation awareness are operationalizations of bounded rationality [109], and expectancy and perceptions are operationalizations of motivation and risk, respectively [110].

The real-life decision-making process deviates from the neoclassical or rational model of decision-making, which assumes perfect information and unlimited absorptive capacity,

time, energy, and other resources—as implied in the framework. The underlying theory for the model is the expectancy-value theory of motivation, which posits that the probability of a specific decision to behave in a particular way (ie, AI-derived decision by a clinician) is dependent on the extent to which the decision maker believes that the specific behavior will elicit an intended outcome (ie, patient safety). The model can be illustrated differently based on the quantitative modeling of future researchers. The framework highlights the shaping factors that are likely to influence clinicians' willingness to use an AI system. The framework captures the way in which the factors influence clinicians' intention to use AI in their clinical workflow. In other words, future researchers can leverage this framework to explore the factors that influence clinicians' cognitive function regarding the use of an AI system and, consecutively, impact the perception of patient safety or risk, trust in AI, and intent to use the technology. Subsequently, the framework also enables us to understand whether and how AI can influence clinical decision-making.

Conflicts of Interest

None declared.

References

1. Higgs J, Jones MA. Clinical decision making and multiple problem spaces. In: Higgs J, Jensen GM, Loftus S, Christensen N, editors. *Clinical Reasoning in the Health Professions*. 3rd edition. Oxford, UK: Butterworth-Heinemann; 2008:3-18.
2. Blake C, Pratt W. Collaborative information synthesis I: a model of information behaviors of scientists in medicine and public health. *J Am Soc Inf Sci* 2006 Nov;57(13):1740-1749 [FREE Full text] [doi: [10.1002/asi.20487](https://doi.org/10.1002/asi.20487)]
3. Reason J. *Human Error*. Cambridge, UK: Cambridge University Press; 1990.
4. Plsek PE, Greenhalgh T. Complexity science: the challenge of complexity in health care. *BMJ* 2001 Sep 15;323(7313):625-628 [FREE Full text] [doi: [10.1136/bmj.323.7313.625](https://doi.org/10.1136/bmj.323.7313.625)] [Medline: [11557716](https://pubmed.ncbi.nlm.nih.gov/11557716/)]
5. Plsek PE, Wilson T. Complexity, leadership, and management in healthcare organisations. *BMJ* 2001 Sep 29;323(7315):746-749 [FREE Full text] [doi: [10.1136/bmj.323.7315.746](https://doi.org/10.1136/bmj.323.7315.746)] [Medline: [11576986](https://pubmed.ncbi.nlm.nih.gov/11576986/)]
6. Patel VL, Zhang J, Yoskowitz NA, Green R, Sayan OR. Translational cognition for decision support in critical care environments: a review. *J Biomed Inform* 2008 Jun;41(3):413-431 [FREE Full text] [doi: [10.1016/j.jbi.2008.01.013](https://doi.org/10.1016/j.jbi.2008.01.013)] [Medline: [18343731](https://pubmed.ncbi.nlm.nih.gov/18343731/)]
7. Gretton C. Trust and transparency in machine learning-based clinical decision support. In: Zhou J, Chen F, editors. *Human and Machine Learning: Visible, Explainable, Trustworthy and Transparent*. Cham, Switzerland: Springer; 2018:279-292.
8. Cypko MA, Stoehr M, Kozniowski M, Druzdzel MJ, Dietz A, Berliner L, et al. Validation workflow for a clinical Bayesian network model in multidisciplinary decision making in head and neck oncology treatment. *Int J Comput Assist Radiol Surg* 2017 Nov;12(11):1959-1970. [doi: [10.1007/s11548-017-1531-7](https://doi.org/10.1007/s11548-017-1531-7)] [Medline: [28204986](https://pubmed.ncbi.nlm.nih.gov/28204986/)]
9. Lanceley A, Savage J, Menon U, Jacobs I. Influences on multidisciplinary team decision-making. *Int J Gynecol Cancer* 2008;18(2):215-222. [doi: [10.1111/j.1525-1438.2007.00991.x](https://doi.org/10.1111/j.1525-1438.2007.00991.x)] [Medline: [17511801](https://pubmed.ncbi.nlm.nih.gov/17511801/)]
10. Wekre LJ, Melby L, Grimsmo A. Early experiences with the multidose drug dispensing system--a matter of trust? *Scand J Prim Health Care* 2011 Mar;29(1):45-50 [FREE Full text] [doi: [10.3109/02813432.2011.554002](https://doi.org/10.3109/02813432.2011.554002)] [Medline: [21323496](https://pubmed.ncbi.nlm.nih.gov/21323496/)]
11. Browne M, Cook P. Inappropriate trust in technology: implications for critical care nurses. *Nurs Crit Care* 2011;16(2):92-98. [doi: [10.1111/j.1478-5153.2010.00407.x](https://doi.org/10.1111/j.1478-5153.2010.00407.x)] [Medline: [21299762](https://pubmed.ncbi.nlm.nih.gov/21299762/)]
12. Hijazi R. The impact of medical devices on patient health: an analysis of patient safety and risk. *J Clin Eng* 2011 Jul;36(3):105-108. [doi: [10.1097/JCE.0b013e3182223c7b2](https://doi.org/10.1097/JCE.0b013e3182223c7b2)]
13. Ho G, Wheatley D, Scialfa CT. Age differences in trust and reliance of a medication management system. *Interact Comput* 2005 Dec;17(6):690-710 [FREE Full text] [doi: [10.1016/j.intcom.2005.09.007](https://doi.org/10.1016/j.intcom.2005.09.007)]
14. Wu JH, Shen WS, Lin LM, Greenes RA, Bates DW. Testing the technology acceptance model for evaluating healthcare professionals' intention to use an adverse event reporting system. *Int J Qual Health Care* 2008 Apr;20(2):123-129. [doi: [10.1093/intqhc/mzm074](https://doi.org/10.1093/intqhc/mzm074)] [Medline: [18222963](https://pubmed.ncbi.nlm.nih.gov/18222963/)]
15. Kleinman Z. Most healthcare apps not up to NHS standards. BBC News. 2021 Feb 16. URL: <https://www.bbc.com/news/technology-56083231> [accessed 2022-05-10]

16. Mnih V, Kavukcuoglu K, Silver D, Graves A, Antonoglou I, Wierstra D. Playing atari with deep reinforcement learning. arXiv Preprint posted online December 19, 2013. [doi: [10.48550/arXiv.1312.5602](https://doi.org/10.48550/arXiv.1312.5602)]
17. Adegboro CO, Choudhury A, Asan O, Kelly MM. Artificial intelligence to improve health outcomes in the NICU and PICU: a systematic review. *Hosp Pediatr* 2022 Jan 01;12(1):93-110. [doi: [10.1542/hpeds.2021-006094](https://doi.org/10.1542/hpeds.2021-006094)] [Medline: [34890453](https://pubmed.ncbi.nlm.nih.gov/34890453/)]
18. Choudhury A, Renjilian E, Asan O. Use of machine learning in geriatric clinical care for chronic diseases: a systematic literature review. *JAMIA Open* 2020 Oct;3(3):459-471 [FREE Full text] [doi: [10.1093/jamiaopen/ooaa034](https://doi.org/10.1093/jamiaopen/ooaa034)] [Medline: [33215079](https://pubmed.ncbi.nlm.nih.gov/33215079/)]
19. Lin S. A clinician's guide to artificial intelligence (AI): why and how primary care should lead the health care AI revolution. *J Am Board Fam Med* 2022;35(1):175-184 [FREE Full text] [doi: [10.3122/jabfm.2022.01.210226](https://doi.org/10.3122/jabfm.2022.01.210226)] [Medline: [35039425](https://pubmed.ncbi.nlm.nih.gov/35039425/)]
20. Martínez-Plumed F, Gómez E, Hernández-Orallo J. Futures of artificial intelligence through technology readiness levels. *Telemat Inform* 2021 May;58:101525 [FREE Full text] [doi: [10.1016/j.tele.2020.101525](https://doi.org/10.1016/j.tele.2020.101525)]
21. Panch T, Mattie H, Celi LA. The "inconvenient truth" about AI in healthcare. *NPJ Digit Med* 2019 Aug 16;2:77 [FREE Full text] [doi: [10.1038/s41746-019-0155-4](https://doi.org/10.1038/s41746-019-0155-4)] [Medline: [31453372](https://pubmed.ncbi.nlm.nih.gov/31453372/)]
22. Rajkomar A, Dean J, Kohane I. Machine learning in medicine. *N Engl J Med* 2019 Apr 04;380(14):1347-1358. [doi: [10.1056/NEJMra1814259](https://doi.org/10.1056/NEJMra1814259)] [Medline: [30943338](https://pubmed.ncbi.nlm.nih.gov/30943338/)]
23. Straub J. In search of technology readiness level (TRL) 10. *Aerosp Sci Technol* 2015 Oct;46:312-320 [FREE Full text] [doi: [10.1016/j.ast.2015.07.007](https://doi.org/10.1016/j.ast.2015.07.007)]
24. Ortega Egea JM, Román González MV. Explaining physicians' acceptance of EHCR systems: an extension of TAM with trust and risk factors. *Comput Human Behav* 2011 Jan;27(1):319-332 [FREE Full text] [doi: [10.1016/j.chb.2010.08.010](https://doi.org/10.1016/j.chb.2010.08.010)]
25. Siegrist M. The influence of trust and perceptions of risks and benefits on the acceptance of gene technology. *Risk Anal* 2000 Apr;20(2):195-203. [doi: [10.1111/0272-4332.202020](https://doi.org/10.1111/0272-4332.202020)] [Medline: [10859780](https://pubmed.ncbi.nlm.nih.gov/10859780/)]
26. Patel VL, Kannampallil TG, Shortliffe EH. Role of cognition in generating and mitigating clinical errors. *BMJ Qual Saf* 2015 Jul;24(7):468-474. [doi: [10.1136/bmjqs-2014-003482](https://doi.org/10.1136/bmjqs-2014-003482)] [Medline: [25935928](https://pubmed.ncbi.nlm.nih.gov/25935928/)]
27. Nsoesie EO. Evaluating artificial intelligence applications in clinical settings. *JAMA Netw Open* 2018 Sep 07;1(5):e182658 [FREE Full text] [doi: [10.1001/jamanetworkopen.2018.2658](https://doi.org/10.1001/jamanetworkopen.2018.2658)] [Medline: [30646173](https://pubmed.ncbi.nlm.nih.gov/30646173/)]
28. van den Bosch K, Bronkhorst A. Human-AI cooperation to benefit military decision making. North Atlantic Treaty Organization - Science and Technology Organization. 2018. URL: <https://www.sto.nato.int/publications/STO%20Meeting%20Proceedings/STO-MP-IST-160/MP-IST-160-S3-1.pdf> [accessed 2022-05-10]
29. Bandaragoda T, Adikari A, Nawaratne R, Nallaperuma D, Luhach AK, Kempitiya T, et al. Artificial intelligence based commuter behaviour profiling framework using Internet of things for real-time decision-making. *Neural Comput & Applic* 2020 Feb 20;32(20):16057-16071 [FREE Full text] [doi: [10.1007/s00521-020-04736-7](https://doi.org/10.1007/s00521-020-04736-7)]
30. Shrestha YR, Ben-Menahem SM, von Krogh G. Organizational decision-making structures in the age of artificial intelligence. *Calif Manag Rev* 2019 Jul 13;61(4):66-83. [doi: [10.1177/0008125619862257](https://doi.org/10.1177/0008125619862257)]
31. Widmer MA, Swanson RC, Zink BJ, Pines JM. Complex systems thinking in emergency medicine: a novel paradigm for a rapidly changing and interconnected health care landscape. *J Eval Clin Pract* 2018 Jun;24(3):629-634. [doi: [10.1111/jep.12862](https://doi.org/10.1111/jep.12862)] [Medline: [29280244](https://pubmed.ncbi.nlm.nih.gov/29280244/)]
32. Qadri YA, Nauman A, Zikria YB, Vasilakos AV, Kim SW. The future of healthcare Internet of Things: a survey of emerging technologies. *IEEE Commun Surv Tutor* 2020 Feb 11;22(2):1121-1167. [doi: [10.1109/comst.2020.2973314](https://doi.org/10.1109/comst.2020.2973314)]
33. Wagstaff K. Robotic Surgery Involved in 144 Deaths in 14 Years. NBC News. 2015 Jul 22. URL: <https://www.nbcnews.com/tech/tech-news/robotic-surgery-linked-144-deaths-2000-n395811> [accessed 2022-06-07]
34. Winfield AF, Jirotko M. Ethical governance is essential to building trust in robotics and artificial intelligence systems. *Philos Trans A Math Phys Eng Sci* 2018 Oct 15;376(2133):20180085 [FREE Full text] [doi: [10.1098/rsta.2018.0085](https://doi.org/10.1098/rsta.2018.0085)] [Medline: [30323000](https://pubmed.ncbi.nlm.nih.gov/30323000/)]
35. Reddy S, Allan S, Coghlan S, Cooper P. A governance model for the application of AI in health care. *J Am Med Inform Assoc* 2020 Mar 01;27(3):491-497 [FREE Full text] [doi: [10.1093/jamia/ocz192](https://doi.org/10.1093/jamia/ocz192)] [Medline: [31682262](https://pubmed.ncbi.nlm.nih.gov/31682262/)]
36. Price 2nd WN, Gerke S, Cohen IG. Potential liability for physicians using artificial intelligence. *JAMA* 2019 Nov 12;322(18):1765-1766. [doi: [10.1001/jama.2019.15064](https://doi.org/10.1001/jama.2019.15064)] [Medline: [31584609](https://pubmed.ncbi.nlm.nih.gov/31584609/)]
37. Carayon P, Schoofs Hundt A, Karsh B, Gurses AP, Alvarado CJ, Smith M, et al. Work system design for patient safety: the SEIPS model. *Qual Saf Health Care* 2006 Dec;15 Suppl 1:i50-i58 [FREE Full text] [doi: [10.1136/qshc.2005.015842](https://doi.org/10.1136/qshc.2005.015842)] [Medline: [17142610](https://pubmed.ncbi.nlm.nih.gov/17142610/)]
38. Donabedian A. Evaluating the quality of medical care. *Milbank Mem Fund Q* 1966 Jul;44(3):166-206. [Medline: [5338568](https://pubmed.ncbi.nlm.nih.gov/5338568/)]
39. Khan OF, Bebb G, Alimohamed NA. Artificial intelligence in medicine: What oncologists need to know about its potential — and its limitations. *Oncol Exch* 2017 Nov;16(4):p8-13.
40. Mendelson EB. Artificial intelligence in breast imaging: potentials and limitations. *AJR Am J Roentgenol* 2019 Feb;212(2):293-299. [doi: [10.2214/AJR.18.20532](https://doi.org/10.2214/AJR.18.20532)] [Medline: [30422715](https://pubmed.ncbi.nlm.nih.gov/30422715/)]
41. Hashimoto DA, Witkowski E, Gao L, Meireles O, Rosman G. Artificial intelligence in anesthesiology: current techniques, clinical applications, and limitations. *Anesthesiology* 2020 Feb;132(2):379-394 [FREE Full text] [doi: [10.1097/ALN.0000000000002960](https://doi.org/10.1097/ALN.0000000000002960)] [Medline: [31939856](https://pubmed.ncbi.nlm.nih.gov/31939856/)]

42. Kappagoda A. The power of systems thinking in medicine. *Med J Aust* 2014 Mar 03;200(4):189. [doi: [10.5694/mja14.c0303](https://doi.org/10.5694/mja14.c0303)] [Medline: [24580507](https://pubmed.ncbi.nlm.nih.gov/24580507/)]
43. Choudhury A, Asan O. Human factors: bridging artificial intelligence and patient safety. *Proc Int Symp Hum Factors Ergon Healthc* 2020 Sep 16;9(1):211-215. [doi: [10.1177/2327857920091007](https://doi.org/10.1177/2327857920091007)]
44. Beaulieu L, Freeman M. Nursing shortcuts can shortcut safety. *Nursing* 2009 Dec;39(12):16-17. [doi: [10.1097/01.NURSE.0000365016.08493.41](https://doi.org/10.1097/01.NURSE.0000365016.08493.41)] [Medline: [19934734](https://pubmed.ncbi.nlm.nih.gov/19934734/)]
45. Abuhav I. ISO 13485:2016: A Complete Guide to Quality Management in the Medical Device Industry. 2nd edition. Boca Raton, FL, USA: CRC Press; May 11, 2016.
46. CFR - Code of Federal Regulations Title 21. U.S. Food & Drug Administration. 2020. URL: <https://www.accessdata.fda.gov/scripts/cdrh/cfdocs/cfCFR/CFRSearch.cfm?fr=820.3> [accessed 2021-03-28]
47. Farhadpoor MR, Soltani G. A study of users' motivation in using information services of regional information Center for Science and Technology. *Lib Philos Pract* 2017:1649.
48. O'Reilly III CA. Variations in decision makers' use of information sources: the impact of quality and accessibility of information. *Acad Manag J* 1982 Dec 01;25(4):756-771 [FREE Full text] [doi: [10.2307/256097](https://doi.org/10.2307/256097)]
49. Liao SH, Fei WC, Chen CC. Knowledge sharing, absorptive capacity, and innovation capability: an empirical study of Taiwan's knowledge-intensive industries. *J Inf Sci* 2007 Jun 1;33(3):340-359 [FREE Full text] [doi: [10.1177/0165551506070739](https://doi.org/10.1177/0165551506070739)]
50. Schmidt T. Absorptive capacity-one size fits all? A firm-level analysis of absorptive capacity for different kinds of knowledge. *Manage Decis Econ* 2009 Jan;31(1):1-18 [FREE Full text] [doi: [10.1002/mde.1423](https://doi.org/10.1002/mde.1423)]
51. Streufert SC. Effects of information relevance on decision making in complex environments. *Mem Cogn* 1973 Sep;1(3):224-228 [FREE Full text] [doi: [10.3758/bf03198100](https://doi.org/10.3758/bf03198100)]
52. Citroen CL. The role of information in strategic decision-making. *Int J Inf Manag* 2011 Dec;31(6):493-501 [FREE Full text] [doi: [10.1016/j.ijinfomgt.2011.02.005](https://doi.org/10.1016/j.ijinfomgt.2011.02.005)]
53. Davis FD, Bagozzi RP, Warshaw PR. User acceptance of computer technology: a comparison of two theoretical models. *Manag Sci* 1989 Aug 1;35(8):982-1003. [doi: [10.1287/mnsc.35.8.982](https://doi.org/10.1287/mnsc.35.8.982)]
54. Todd P, Benbasat I. The use of information in decision making: an experimental investigation of the impact of computer-based decision aids. *MIS Q* 1992 Sep;16(3):373-393 [FREE Full text] [doi: [10.2307/249534](https://doi.org/10.2307/249534)]
55. Snead Jr KC, Harrell AM. An application of expectancy theory to explain a manager's intention to use a decision support system. *Decis Sci* 1994 Jul;25(4):499-510 [FREE Full text] [doi: [10.1111/j.1540-5915.1994.tb01857.x](https://doi.org/10.1111/j.1540-5915.1994.tb01857.x)]
56. Shi G, Ma Z, Feng J, Zhu F, Bai X, Gui B. The impact of knowledge transfer performance on the artificial intelligence industry innovation network: an empirical study of Chinese firms. *PLoS One* 2020 May 18;15(5):e0232658 [FREE Full text] [doi: [10.1371/journal.pone.0232658](https://doi.org/10.1371/journal.pone.0232658)] [Medline: [32421743](https://pubmed.ncbi.nlm.nih.gov/32421743/)]
57. Prevedello LM, Halabi SS, Shih G, Wu CC, Kohli MD, Chokshi FH, et al. Challenges related to artificial intelligence research in medical imaging and the importance of image analysis competitions. *Radiol Artif Intell* 2019 Jan;1(1):e180031 [FREE Full text] [doi: [10.1148/ryai.2019180031](https://doi.org/10.1148/ryai.2019180031)] [Medline: [33937783](https://pubmed.ncbi.nlm.nih.gov/33937783/)]
58. Pomeroy JC. Artificial intelligence and human decision making. *Eur J Oper Res* 1997 May 16;99(1):3-25 [FREE Full text] [doi: [10.1016/s0377-2217\(96\)00378-5](https://doi.org/10.1016/s0377-2217(96)00378-5)]
59. Chan KS, Zary N. Applications and challenges of implementing artificial intelligence in medical education: integrative review. *JMIR Med Educ* 2019 Jun 15;5(1):e13930 [FREE Full text] [doi: [10.2196/13930](https://doi.org/10.2196/13930)] [Medline: [31199295](https://pubmed.ncbi.nlm.nih.gov/31199295/)]
60. Sohn K, Kwon O. Technology acceptance theories and factors influencing artificial intelligence-based intelligent products. *Telemat Inform* 2020 Apr;47:101324 [FREE Full text] [doi: [10.1016/j.tele.2019.101324](https://doi.org/10.1016/j.tele.2019.101324)]
61. Lowry PB, Zhang D, Zhou L, Fu X. Effects of culture, social presence, and group composition on trust in technology-supported decision-making groups. *Inf Syst J* 2010 May;20(3):297-315 [FREE Full text] [doi: [10.1111/j.1365-2575.2009.00334.x](https://doi.org/10.1111/j.1365-2575.2009.00334.x)]
62. Johnson DS. Achieving customer value from electronic channels through identity commitment, calculative commitment, and trust in technology. *J Interact Mark* 2007 Aug 01;21(4):2-22 [FREE Full text] [doi: [10.1002/dir.20091](https://doi.org/10.1002/dir.20091)]
63. Downer J. Trust and technology: the social foundations of aviation regulation. *Br J Sociol* 2010 Mar;61(1):83-106. [doi: [10.1111/j.1468-4446.2009.01303.x](https://doi.org/10.1111/j.1468-4446.2009.01303.x)] [Medline: [20377598](https://pubmed.ncbi.nlm.nih.gov/20377598/)]
64. McKnight DH, Carter M, Thatcher JB, Clay PF. Trust in a specific technology: an investigation of its components and measures. *ACM Trans Manage Inf Syst* 2011 Jun 01;2(2):1-25. [doi: [10.1145/1985347.1985353](https://doi.org/10.1145/1985347.1985353)]
65. Lankton NK, McKnight DH. What does it mean to trust Facebook?: examining technology and interpersonal trust beliefs. *SIGMIS Database* 2011 May 24;42(2):32-54 [FREE Full text] [doi: [10.1145/1989098.1989101](https://doi.org/10.1145/1989098.1989101)]
66. Oliver RL. Effect of expectation and disconfirmation on postexposure product evaluations: an alternative interpretation. *J Appl Psychol* 1977;62(4):480-486. [doi: [10.1037/0021-9010.62.4.480](https://doi.org/10.1037/0021-9010.62.4.480)]
67. Li X, Hess TJ, Valacich JS. Why do we trust new technology? A study of initial trust formation with organizational information systems. *J Strateg Inf Syst* 2008 Mar;17(1):39-71 [FREE Full text] [doi: [10.1016/j.jsis.2008.01.001](https://doi.org/10.1016/j.jsis.2008.01.001)]
68. Montague EN, Kleiner BM, Winchester III WW. Empirically understanding trust in medical technology. *Int J Ind Ergon* 2009 Jul;39(4):628-634 [FREE Full text] [doi: [10.1016/j.ergon.2009.01.004](https://doi.org/10.1016/j.ergon.2009.01.004)]
69. Hurlburt G. How much to trust artificial intelligence? *IT Prof* 2017 Aug 17;19(4):7-11. [doi: [10.1109/mitp.2017.3051326](https://doi.org/10.1109/mitp.2017.3051326)]

70. Hengstler M, Enkel E, Duelli S. Applied artificial intelligence and trust—the case of autonomous vehicles and medical assistance devices. *Technol Forecast Soc Change* 2016 Apr;105:105-120 [FREE Full text] [doi: [10.1016/j.techfore.2015.12.014](https://doi.org/10.1016/j.techfore.2015.12.014)]
71. Glikson E, Woolley AW. Human trust in artificial intelligence: review of empirical research. *Acad Manag Ann* 2020 Jul;14(2):627-660. [doi: [10.5465/annals.2018.0057](https://doi.org/10.5465/annals.2018.0057)]
72. Kerasidou A. Artificial intelligence and the ongoing need for empathy, compassion and trust in healthcare. *Bull World Health Organ* 2020 Apr 01;98(4):245-250 [FREE Full text] [doi: [10.2471/BLT.19.237198](https://doi.org/10.2471/BLT.19.237198)] [Medline: [32284647](https://pubmed.ncbi.nlm.nih.gov/32284647/)]
73. Jiang F, Jiang Y, Zhi H, Dong Y, Li H, Ma S, et al. Artificial intelligence in healthcare: past, present and future. *Stroke Vasc Neurol* 2017 Jun 21;2(4):230-243 [FREE Full text] [doi: [10.1136/svn-2017-000101](https://doi.org/10.1136/svn-2017-000101)] [Medline: [29507784](https://pubmed.ncbi.nlm.nih.gov/29507784/)]
74. Banerjee I, Li K, Seneviratne M, Ferrari M, Seto T, Brooks JD, et al. Weakly supervised natural language processing for assessing patient-centered outcome following prostate cancer treatment. *JAMIA Open* 2019 Apr;2(1):150-159 [FREE Full text] [doi: [10.1093/jamiaopen/ooy057](https://doi.org/10.1093/jamiaopen/ooy057)] [Medline: [31032481](https://pubmed.ncbi.nlm.nih.gov/31032481/)]
75. Ciervo J, Shen SC, Stallcup K, Thomas A, Farnum MA, Lobanov VS, et al. A new risk and issue management system to improve productivity, quality, and compliance in clinical trials. *JAMIA Open* 2019 Jul;2(2):216-221 [FREE Full text] [doi: [10.1093/jamiaopen/ooz006](https://doi.org/10.1093/jamiaopen/ooz006)] [Medline: [31984356](https://pubmed.ncbi.nlm.nih.gov/31984356/)]
76. Ronquillo JG, Erik Winterholler J, Cwikla K, Szymanski R, Levy C. Health IT, hacking, and cybersecurity: national trends in data breaches of protected health information. *JAMIA Open* 2018 Jul;1(1):15-19 [FREE Full text] [doi: [10.1093/jamiaopen/ooy019](https://doi.org/10.1093/jamiaopen/ooy019)] [Medline: [31984315](https://pubmed.ncbi.nlm.nih.gov/31984315/)]
77. Dalal AK, Fuller T, Garabedian P, Ergai A, Balint C, Bates DW, et al. Systems engineering and human factors support of a system of novel EHR-integrated tools to prevent harm in the hospital. *J Am Med Inform Assoc* 2019 Jun 01;26(6):553-560 [FREE Full text] [doi: [10.1093/jamia/ocz002](https://doi.org/10.1093/jamia/ocz002)] [Medline: [30903660](https://pubmed.ncbi.nlm.nih.gov/30903660/)]
78. Sanchez-Morillo D, Fernandez-Granero MA, Leon-Jimenez A. Use of predictive algorithms in-home monitoring of chronic obstructive pulmonary disease and asthma: a systematic review. *Chron Respir Dis* 2016 Aug;13(3):264-283 [FREE Full text] [doi: [10.1177/1479972316642365](https://doi.org/10.1177/1479972316642365)] [Medline: [27097638](https://pubmed.ncbi.nlm.nih.gov/27097638/)]
79. Pellegrini E, Ballerini L, Hernandez MD, Chappell FM, González-Castro V, Anblagan D, et al. Machine learning of neuroimaging for assisted diagnosis of cognitive impairment and dementia: a systematic review. *Alzheimers Dement (Amst)* 2018 Aug 11;10:519-535 [FREE Full text] [doi: [10.1016/j.dadm.2018.07.004](https://doi.org/10.1016/j.dadm.2018.07.004)] [Medline: [30364671](https://pubmed.ncbi.nlm.nih.gov/30364671/)]
80. Safdar S, Zafar S, Zafar N, Khan NF. Machine learning based decision support systems (DSS) for heart disease diagnosis: a review. *Artif Intell Rev* 2018 Dec;50(4):597-623. [doi: [10.1007/s10462-017-9552-8](https://doi.org/10.1007/s10462-017-9552-8)]
81. Dallora AL, Eivazzadeh S, Mendes E, Berglund J, Anderberg P. Machine learning and microsimulation techniques on the prognosis of dementia: a systematic literature review. *PLoS One* 2017 Jun 29;12(6):e0179804 [FREE Full text] [doi: [10.1371/journal.pone.0179804](https://doi.org/10.1371/journal.pone.0179804)] [Medline: [28662070](https://pubmed.ncbi.nlm.nih.gov/28662070/)]
82. Liu X, Faes L, Kale AU, Wagner SK, Fu DJ, Bruynseels A, et al. A comparison of deep learning performance against health-care professionals in detecting diseases from medical imaging: a systematic review and meta-analysis. *Lancet Digit Health* 2019 Oct;1(6):e271-e297 [FREE Full text] [doi: [10.1016/S2589-7500\(19\)30123-2](https://doi.org/10.1016/S2589-7500(19)30123-2)] [Medline: [33323251](https://pubmed.ncbi.nlm.nih.gov/33323251/)]
83. Chen M, Hao Y, Hwang K, Wang L, Wang L. Disease prediction by machine learning over big data from healthcare communities. *IEEE Access* 2017 Apr 26;5:8869-8879. [doi: [10.1109/ACCESS.2017.2694446](https://doi.org/10.1109/ACCESS.2017.2694446)]
84. Yeh DH, Tam S, Fung K, MacNeil SD, Yoo J, Winquist E, et al. Transoral robotic surgery vs. radiotherapy for management of oropharyngeal squamous cell carcinoma - a systematic review of the literature. *Eur J Surg Oncol* 2015 Dec;41(12):1603-1614. [doi: [10.1016/j.ejso.2015.09.007](https://doi.org/10.1016/j.ejso.2015.09.007)] [Medline: [26461255](https://pubmed.ncbi.nlm.nih.gov/26461255/)]
85. Ficarra V, Novara G, Rosen RC, Artibani W, Carroll PR, Costello A, et al. Systematic review and meta-analysis of studies reporting urinary continence recovery after robot-assisted radical prostatectomy. *Eur Urol* 2012 Sep;62(3):405-417. [doi: [10.1016/j.eururo.2012.05.045](https://doi.org/10.1016/j.eururo.2012.05.045)] [Medline: [22749852](https://pubmed.ncbi.nlm.nih.gov/22749852/)]
86. Douthwaite SA, Franklin JH, Palma DA, Fung K, Yoo J, Nichols AC. The role of transoral robotic surgery in the management of oropharyngeal cancer: a review of the literature. *ISRN Oncol* 2012;2012:945162 [FREE Full text] [doi: [10.5402/2012/945162](https://doi.org/10.5402/2012/945162)] [Medline: [22606380](https://pubmed.ncbi.nlm.nih.gov/22606380/)]
87. Karthik K, Colegate-Stone T, Dasgupta P, Tavakkolizadeh A, Sinha J. Robotic surgery in trauma and orthopaedics: a systematic review. *Bone Joint J* 2015 Mar;97-B(3):292-299. [doi: [10.1302/0301-620X.97B3.35107](https://doi.org/10.1302/0301-620X.97B3.35107)] [Medline: [25737510](https://pubmed.ncbi.nlm.nih.gov/25737510/)]
88. Klock M, Kang H, Gong Y. Scoring patient fall reports using quality rubric and machine learning. *Stud Health Technol Inform* 2019 Aug 21;264:639-643. [doi: [10.3233/SHTI190301](https://doi.org/10.3233/SHTI190301)] [Medline: [31438002](https://pubmed.ncbi.nlm.nih.gov/31438002/)]
89. Wang E, Kang H, Gong Y. Generating a health information technology event database from FDA MAUDE reports. *Stud Health Technol Inform* 2019 Aug 21;264:883-887. [doi: [10.3233/SHTI190350](https://doi.org/10.3233/SHTI190350)] [Medline: [31438051](https://pubmed.ncbi.nlm.nih.gov/31438051/)]
90. Zhou S, Kang H, Yao B, Gong Y. An automated pipeline for analyzing medication event reports in clinical settings. *BMC Med Inform Decis Mak* 2018 Dec 07;18(Suppl 5):113 [FREE Full text] [doi: [10.1186/s12911-018-0687-6](https://doi.org/10.1186/s12911-018-0687-6)] [Medline: [30526590](https://pubmed.ncbi.nlm.nih.gov/30526590/)]
91. Woolf SH, Chan EC, Harris R, Sheridan SL, Braddock 3rd CH, Kaplan RM, et al. Promoting informed choice: transforming health care to dispense knowledge for decision making. *Ann Intern Med* 2005 Aug 16;143(4):293-300. [doi: [10.7326/0003-4819-143-4-200508160-00010](https://doi.org/10.7326/0003-4819-143-4-200508160-00010)] [Medline: [16103473](https://pubmed.ncbi.nlm.nih.gov/16103473/)]

92. L egar  F, Adekpedjou R, Stacey D, Turcotte S, Kryworuchko J, Graham ID, et al. Interventions for increasing the use of shared decision making by healthcare professionals. *Cochrane Database Syst Rev* 2018 Jul 19;7(7):CD006732 [FREE Full text] [doi: [10.1002/14651858.CD006732.pub4](https://doi.org/10.1002/14651858.CD006732.pub4)] [Medline: [30025154](https://pubmed.ncbi.nlm.nih.gov/30025154/)]
93. Edwards M, Davies M, Edwards A. What are the external influences on information exchange and shared decision-making in healthcare consultations: a meta-synthesis of the literature. *Patient Educ Couns* 2009 Apr;75(1):37-52. [doi: [10.1016/j.pec.2008.09.025](https://doi.org/10.1016/j.pec.2008.09.025)] [Medline: [19036550](https://pubmed.ncbi.nlm.nih.gov/19036550/)]
94. Bindels J, Ramaekers B, Ramos IC, Mohseninejad L, Knies S, Grutters J, et al. Use of value of information in healthcare decision making: exploring multiple perspectives. *Pharmacoeconomics* 2016 Mar;34(3):315-322 [FREE Full text] [doi: [10.1007/s40273-015-0346-z](https://doi.org/10.1007/s40273-015-0346-z)] [Medline: [26578403](https://pubmed.ncbi.nlm.nih.gov/26578403/)]
95. Tuffaha HW, Gordon LG, Scuffham PA. Value of information analysis in oncology: the value of evidence and evidence of value. *J Oncol Pract* 2014 Mar;10(2):e55-e62. [doi: [10.1200/JOP.2013.001108](https://doi.org/10.1200/JOP.2013.001108)] [Medline: [24194511](https://pubmed.ncbi.nlm.nih.gov/24194511/)]
96. Smith K, Hancock PA. Situation awareness is adaptive, externally directed consciousness. *Hum Factors* 1995 Mar 1;37(1):137-148. [doi: [10.1518/001872095779049444](https://doi.org/10.1518/001872095779049444)]
97. Bedny G, Meister D. Theory of activity and situation awareness. *Int J Cogn Ergon* 1999 Jan;3(1):63-72. [doi: [10.1207/s15327566ijce0301_5](https://doi.org/10.1207/s15327566ijce0301_5)]
98. Nygren TE. Psychometric properties of subjective workload measurement techniques: implications for their use in the assessment of perceived mental workload. *Hum Factors* 1991 Feb 1;33(1):17-33. [doi: [10.1177/001872089103300102](https://doi.org/10.1177/001872089103300102)]
99. Lyddiardc J, Briggs J, Coffey M, Berridge J. A workload measurement. *Applied Clinical Trials*. 2010 Aug 2. URL: <https://www.appliedclinicaltrials.com/view/workload-measurement> [accessed 2022-05-10]
100. Wang D, Weisz JD, Muller M, Ram P, Geyer W, Dugan C, et al. Human-AI collaboration in data science: exploring data scientists' perceptions of automated AI. *Proc ACM Hum Comput Interact* 2019 Nov 07;3(CSCW):1-24. [doi: [10.1145/3359313](https://doi.org/10.1145/3359313)]
101. Salminen J, Liu YH, Seng n S, Santos JM, Jung SG, Jansen BJ. The effect of numerical and textual information on visual engagement and perceptions of AI-driven persona interfaces. In: *Proceedings of the 25th International Conference on Intelligent User Interfaces*. 2020 Mar Presented at: IUI '20; March 17-20, 2020; Cagliari, Italy p. 357-368. [doi: [10.1145/3377325.3377492](https://doi.org/10.1145/3377325.3377492)]
102. Arnold M, Bellamy RK, Hind M, Houde S, Mehta S, Mojsilovic A, et al. FactSheets: increasing trust in AI services through supplier's declarations of conformity. *IBM J Res Dev* 2019 Sep 18;63(4/5):6:1-613. [doi: [10.1147/jrd.2019.2942288](https://doi.org/10.1147/jrd.2019.2942288)]
103. Lee SH, Phan PH, Dorman T, Weaver SJ, Pronovost PJ. Handoffs, safety culture, and practices: evidence from the hospital survey on patient safety culture. *BMC Health Serv Res* 2016 Jul 12;16:254 [FREE Full text] [doi: [10.1186/s12913-016-1502-7](https://doi.org/10.1186/s12913-016-1502-7)] [Medline: [27405226](https://pubmed.ncbi.nlm.nih.gov/27405226/)]
104. Dang YM, Zhang YG, Brown SA, Chen H. Examining the impacts of mental workload and task-technology fit on user acceptance of the social media search system. *Inf Syst Front* 2018 Sep 8;22(3):697-718. [doi: [10.1007/s10796-018-9879-y](https://doi.org/10.1007/s10796-018-9879-y)]
105. Chao CM. Factors determining the behavioral intention to use mobile learning: an application and extension of the UTAUT model. *Front Psychol* 2019 Jul 16;10:1652 [FREE Full text] [doi: [10.3389/fpsyg.2019.01652](https://doi.org/10.3389/fpsyg.2019.01652)] [Medline: [31379679](https://pubmed.ncbi.nlm.nih.gov/31379679/)]
106. Ghazizadeh M, Peng Y, Lee JD, Boyle LN. Augmenting the technology acceptance model with trust: commercial drivers' attitudes towards monitoring and feedback. *Proc Hum Factors Ergon Soc Annu Meet* 2012 Sep 01;56(1):2286-2290. [doi: [10.1177/1071181312561481](https://doi.org/10.1177/1071181312561481)]
107. Matthews MD, Beal SA. Assessing situation awareness in field training exercises. U.S. Army Research Institute for the Behavioral and Social Sciences. 2002 Sep. URL: <https://apps.dtic.mil/sti/pdfs/ADA408560.pdf> [accessed 2022-05-10]
108. Jane GV. *Human Performance and Situation Awareness Measures*. Boca Raton, FL, USA: CRC Press; 2019.
109. Bazerman MH, Sezer O. Bounded awareness: implications for ethical decision making. *Organ Behav Hum Decis Process* 2016 Sep;136:95-105 [FREE Full text] [doi: [10.1016/j.obhdp.2015.11.004](https://doi.org/10.1016/j.obhdp.2015.11.004)]
110. Baumann MR, Bonner BL. An expectancy theory approach to group coordination: expertise, task features, and member behavior. *J Behav Dec Making* 2017 Apr 01;30(2):407-419 [FREE Full text] [doi: [10.1002/bdm.1954](https://doi.org/10.1002/bdm.1954)]

Abbreviations

- AI:** artificial intelligence
- HFE:** human factors and ergonomics
- HIT:** health information technology
- SEIPS:** Safety Engineering Initiative for Patient Safety
- TRL:** technology readiness level

Edited by A Kushniruk; submitted 03.12.21; peer-reviewed by G Maddern, OS Liang; comments to author 06.03.22; revised version received 26.03.22; accepted 20.05.22; published 21.06.22.

Please cite as:

Choudhury A

Toward an Ecologically Valid Conceptual Framework for the Use of Artificial Intelligence in Clinical Settings: Need for Systems Thinking, Accountability, Decision-making, Trust, and Patient Safety Considerations in Safeguarding the Technology and Clinicians
JMIR Hum Factors 2022;9(2):e35421

URL: <https://humanfactors.jmir.org/2022/2/e35421>

doi: [10.2196/35421](https://doi.org/10.2196/35421)

PMID: [35727615](https://pubmed.ncbi.nlm.nih.gov/35727615/)

©Avishek Choudhury. Originally published in JMIR Human Factors (<https://humanfactors.jmir.org>), 21.06.2022. This is an open-access article distributed under the terms of the Creative Commons Attribution License (<https://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Human Factors, is properly cited. The complete bibliographic information, a link to the original publication on <https://humanfactors.jmir.org>, as well as this copyright and license information must be included.

Original Paper

A Mobile App for Advance Care Planning and Advance Directives (Accordons-nous): Development and Usability Study

Céline Schöpfer¹, MPhil; Frederic Ehrler², PhD; Antoine Berger², MA; Catherine Bollondi Pauly³, BSc; Laurence Buytaert⁴, MD; Camille De La Serna⁴, MD; Florence Hartheiser¹, MD, PhD; Thomas Fassier^{5,6}, MD, PhD; Christine Clavien¹, PhD

¹Institute for Ethics, History, and the Humanities, University Medical Center, University of Geneva, Geneva, Switzerland

²Department of Medical Information Sciences, University Hospitals of Geneva, Geneva, Switzerland

³Direction of Care, University Hospitals of Geneva, Geneva, Switzerland

⁴University Medical Center, University of Geneva, Geneva, Switzerland

⁵Division of Internal Medicine for the Aged, University Hospitals of Geneva, Geneva, Switzerland

⁶Interprofessional Simulation Center, Faculty of Medicine, University of Geneva, Geneva, Switzerland

Corresponding Author:

Céline Schöpfer, MPhil

Institute for Ethics, History, and the Humanities

University Medical Center

University of Geneva

CMU/1 rue Michel Servet

Geneva, 1211

Switzerland

Phone: 41 22 379 46 00

Email: celine.schopfer@etu.unige.ch

Abstract

Background: Advance care planning, including advance directives, is an important tool that allows patients to express their preferences for care if they are no longer able to express themselves. We developed *Accordons-nous*, a smartphone app that informs patients about advance care planning and advance directives, facilitates communication on these sensitive topics, and helps patients express their values and preferences for care.

Objective: The first objective of this study is to conduct a usability test of this app. The second objective is to collect users' critical opinions on the usability and relevance of the tool.

Methods: We conducted a usability test by means of a think-aloud method, asking 10 representative patients to complete 7 browsing tasks. We double coded the filmed sessions to obtain descriptive data on task completion (with or without help), time spent, number of clicks, and the types of problems encountered. We assessed the severity of the problems encountered and identified the modifications needed to address these problems. We evaluated the readability of the app using *Scolarius*, a French equivalent of the Flesch Reading Ease test. By means of a posttest questionnaire, we asked participants to assess the app's usability (System Usability Scale), relevance (Mobile App Rating Scale, section F), and whether they would recommend the app to the target groups: patients, health professionals, and patients' caring relatives.

Results: Participants completed the 7 think-aloud tasks in 80% (56/70) of the cases without any help from the experimenter, in 16% (11/70) of the cases with some help, and failed in 4% (3/70) of the cases. The analysis of failures and difficulties encountered revealed a series of major usability problems that could be addressed with minor modifications to the app. *Accordons-nous* obtained high scores on readability (overall score of 87.4 on *Scolarius* test, corresponding to elementary school level), usability (85.3/100 on System Usability Scale test), relevance (4.3/5 on the Mobile App Rating Scale, section F), and overall subjective endorsement on 3 *I would recommend* questions (4.7/5).

Conclusions: This usability test helped us make the final changes to our app before its official launch.

(*JMIR Hum Factors* 2022;9(2):e34626) doi:[10.2196/34626](https://doi.org/10.2196/34626)

KEYWORDS

usability; mobile apps; advance directives; advance care planning; mHealth; mobile health; palliative care; mobile phone

Introduction

Background

Medical progress increases the availability of treatment options and life-sustaining possibilities. Consequently, it also increases the necessity to make choices between medical options that have different impacts on patients' future lives. When patients are not capable of making those decisions, health care professionals and surrogates are requested to choose, even if they have no clue about the patient's values and priorities. To avoid such distressful and suboptimal situations, early advance care planning (ACP) is increasingly being recognized as an important condition for adequate treatment [1,2]. ACP is "the ability to enable individuals to define goals and preferences for future medical treatment and care, to discuss these goals and preferences with family and health care providers, and to record and review these preferences if appropriate" [3]. An ACP procedure may result in written advance directives (AD); however, for such document to be of any use, it remains important that it is precise and that patients discuss its content with their surrogate decision-maker and update it regularly [4].

Despite the acknowledged importance of ACP, few patients discuss their fears, priorities, and the type of care they would like to receive with their families and professional caregivers [1,5,6]. This is partly because of a lack of knowledge about ACP, a lack of recognition that ACP is relevant for them, and a lack of assistance in this psychologically heavy process [6-8]. Studies also indicate that even health care professionals find it difficult to initiate such discussions [9-11].

ACP can be seen as a process of behavior change involving steps, including awareness and knowledge acquisition, thinking and commitment to act (eg, talking to someone and writing AD), action, and regular updates [12,13]. It has been suggested that patient-centered, computer-based infographics could enhance interest, understanding, recalling, contemplation, and actual sharing of decisions related to ACP [14]. They may assist in the ACP process and save clinicians' time, as patients may be able to obtain relevant information and input to start the ACP process in a family context. Free web-based ACP tools have been developed, with promising preliminary results [15-20]. Mobile apps for ACP are available in some countries. Their potential is recognized; however, existing products have limited features and operate mostly in English only [21,22]. We did not find comparable tools available in French. A recent review

concluded that overall, mobile apps provide insufficient content and features: they are helpful to users who are ready to complete AD rather than those seeking to learn about the ACP process. Moreover, they are poor in terms of design quality, layout, and functionality. Only one app (in English) was assessed as "fairly easy to read" [21].

Accordons-nous

To support the ACP process for French-speaking Swiss residents, our interprofessional research team comprising ethicists, physicians, nurses, patients as partners, information technology professionals developed a free and easy-to-use solution, an app called *Accordons-nous (Let's agree)*.

We developed the content and structure of the app between January 2019 and July 2021 using mixed methods—a Delphi procedure and multiple user tests involving laypeople and patients as partners. [Multimedia Appendix 1 \[23-25\]](#) provides the details about the process of making the app.

Accordons-nous aims to address known barriers to ACP: lack of awareness; insufficient health literacy; difficulty in starting the discussion; identifying one's own preferences, values, and goals of care; and writing personalized and comprehensive AD and updating them regularly within a smooth process.

The app is adapted to the local context of French-speaking Switzerland, designed to be used by patients and their family caregivers independently at home or to support ACP discussions with health professionals. Its content is adapted to the health literacy of ordinary patients. It includes (1) content to make users aware of the importance of anticipating health-related decisions and motivate them to engage in their ACP; (2) essential information on the ACP procedure, legal issues, definitions of technical terms in a format, and vocabulary accessible to common users; (3) discussion prompts to help users engage in ACP discussions with their family, friends, and professional care providers; (4) a card game to help patients clarify their values and priorities in life and reflect on the conditions that would make life not worth living; and (5) a detailed and comprehensive advance directive form that can be tailored to individual life and health situations.

The app, developed in Angular 11 [26], is structured in 3 main sections ([Figures 1-3](#))—*Je m'informe* (I get informed), *J'en parle* (I talk about it), and *J'écris* (I write)—accessible by clicking on a menu at the bottom of the screen. Each section contains several pages accessible via a drop-down menu.

Figure 1. Structure of the section I get information containing subpages accessible through a drop-down menu. ACP: advance care planning; AD: advance directives.

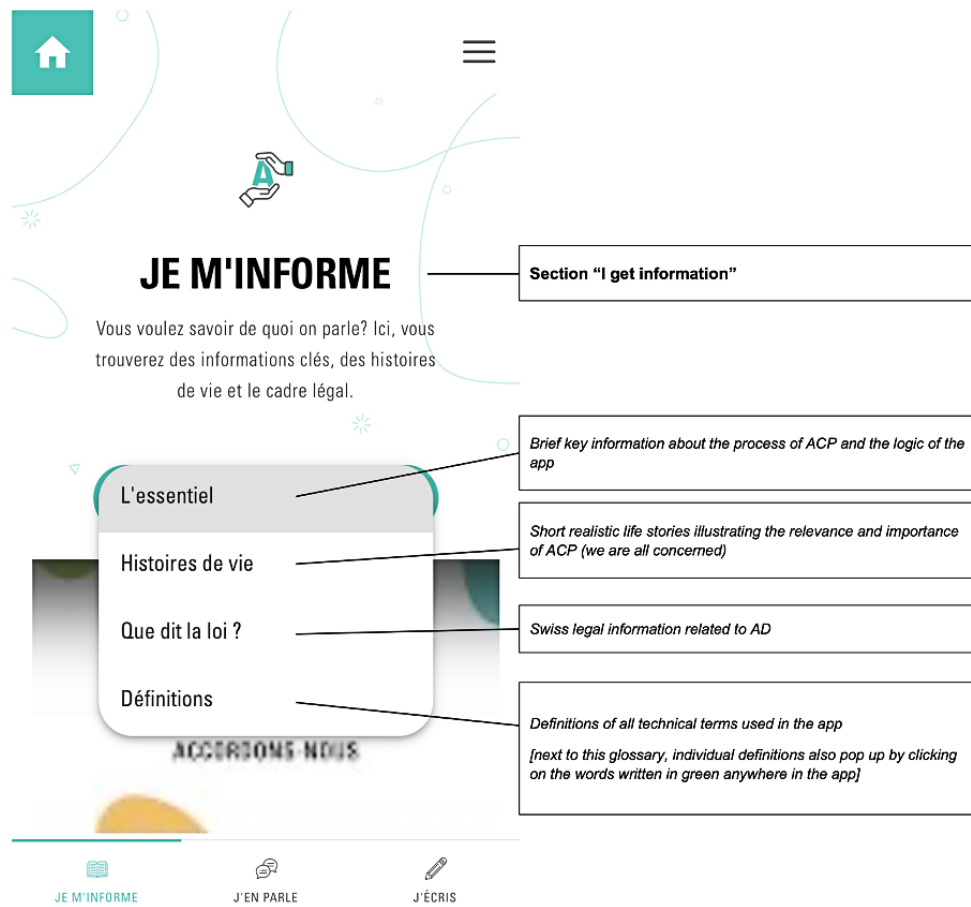
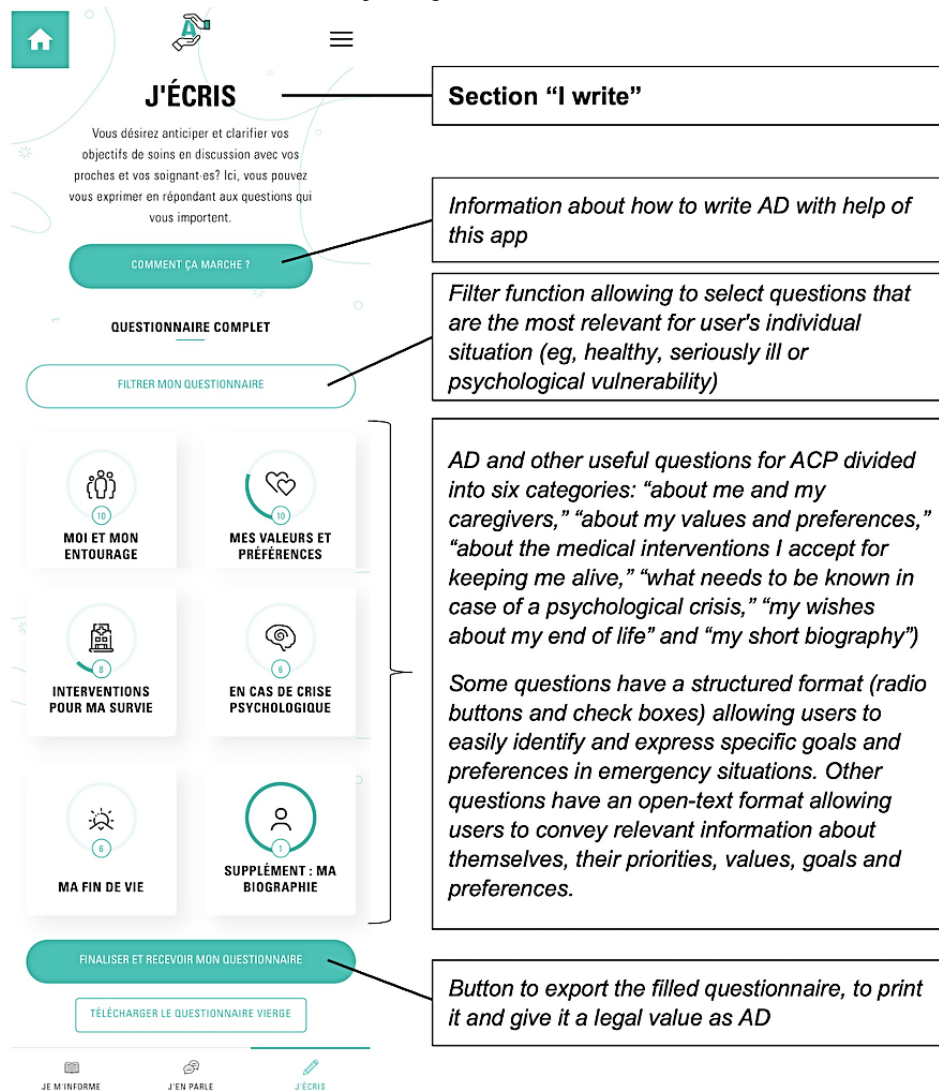


Figure 2. Structure of the section I talk about it with the different pages accessible through a drop-down menu. ACP: advance care planning.



Figure 3. Structure of the section I write. ACP: advance care planning; AD: advance directives.



Study Objectives

The primary objective of this study is to measure whether *Accordons-nous* is easy to use and understand by common users. The secondary objectives are to identify the usability problems encountered by common users and to collect their critical opinions on the relevance of the tool—do they think that the app helps to engage in ACP procedures, and would they recommend the app to its target populations?

Methods

Participants

Nielsen and Landauer [27] reported that 98% of usability problems could be detected through feedback from 10 users. On the basis of this, we aimed to obtain at least 10 complete responses. We asked the *Patients as Partners Project* at the Geneva University Hospitals (Hopitaux Universitaires de Genève [HUG]) to recruit patients fulfilling our criteria. Participation was free and voluntary. The patients received and signed informed consent forms before the study.

We used the following inclusion criteria:

1. Fluency in written and oral French (the app is available only in French)
2. Comfortable with using a mobile phone or tablet
3. Demonstrates a minimal understanding of the concept of AD (to avoid a heavy cognitive load; ordinary users will have more time than our test participants to explore the app and be acquainted with the topic)
4. Did not participate in the development of *Accordons-nous* (to obtain unbiased responses)

Being incapable of understanding the meaning of the questions and tasks required for the study after 3 iterations of the task instructions was the only exclusion criterion.

Ethical Considerations

We submitted our research protocol to the Geneva Commission Cantonale d'Éthique de la Recherche (Req-ID 2020-01397), who decided that the project does not fall within the scope of the Human Research Ordinance and therefore alleviated the need for ethical approval (decision date 12.1.2020). This study was conducted in accordance with the Declaration of Helsinki. Willing participants were informed of the study aims and procedures and signed an informed consent form. During the data collection and analysis phase, only three authors (CS, FE,

and CC) had access to the recorded interviews. The recordings were destroyed after analysis. No identifying information was kept in the remaining work files.

Study Intervention

We conducted a usability test [28] using a think-aloud method [29] comprising 7 tasks. Moreover, we collected quantitative descriptive data using a questionnaire to obtain information about participants' endorsement of the app and their evaluation of the app's usability and relevance for ACP.

We conducted the usability test sessions on the web using the Zoom video conferencing platform. Ahead of the test day, participants received the information and consent sheet and the instructions on how to log on to Zoom by email. Once connected, the experimenter (CS) recalled the procedure as described in the information and consent sheet, answered the participants' questions, and asked for verbal (recorded on Zoom) or written (to be sent after the test) consent. The experimenter also expressed her commitment to respect the study protocol. Verbal consent was recorded separately to protect participants' anonymity during the analysis stage.

Next, the experimenter instructed participants to install the test version of the *Concerto HUG* app that contains the module *Accordons-nous* on their personal smartphone or tablet. The experimenter then instructed the participants to connect their phone or tablet to the same Zoom conversation and share their phone or tablet screen. In this way, while discussing with the experimenter in front of their computer, participants could simultaneously navigate the app via their smartphone or tablet, and the experimenter could see how participants did so (owing to the screen sharing function). Precautions were taken beforehand to avoid personal notifications appearing during the screen sharing.

The experimenter then started the actual test. To ensure that participants fulfilled the inclusion criteria, she began by asking some demographic and technical questions (age, gender, and type of phone or tablet used), as well as participants' understanding of the topic of AD (Multimedia Appendix 1; general questions).

Next, to ensure that participants did not need to express their personal choices and preferences regarding ACP, the experimenter asked the participants to put themselves in the following fictional scenario:

You are a 75-year-old man/woman; you have always been in good health, but you recently started

*experiencing a heart problem that required hospitalization. Fortunately, you recovered fully, but this experience made you think a lot. Now you are asking yourself, "What if another major health accident occurs, during which I lose my capacity to make decisions? How would health professionals know about what matters to me?" You know that it is possible to express your preferences in a document called "advance directives," but you do not know how to do this. You share your concerns with a friend, who tells you about the application *Accordons-nous* and you decide to use it to write your directives. You have just downloaded the application and explore it for the first time.*

Participants were then successively given 7 tasks to complete in the *Accordons-nous* module (Table 1). The tasks were designed to test users' ease while completing the main tasks that a common user of the app is meant to handle and identify the strengths and weaknesses of the app.

While completing the tasks, participants were asked to express all their thoughts verbally (*think-aloud*), as if they were speaking to themselves. If participants forgot to share their reasoning aloud, the examiner reminded them to share their thoughts with her after 15 seconds. Otherwise, the examiner would not intervene. If the task was too complex and the participant felt lost for more than a minute or decided to give up, the examiner offered assistance by providing hints to help the participant move forward in solving the task. When required, the experimenter could provide four types of help: (1) information (eg, explanation of how the participant arrived where they are or why the module includes certain content), (2) motivating question (eg, "Right, and where would you go now?"), (3) answer (if the participant asked a question related to the exploration), and (4) guidance (eg, if the participant lingered too long on an irrelevant page, tell him or her what the next step would help find the answer).

At the end of each task, we asked participants, "Does the path to fulfil the task seem logical to you?" and "Do you have any suggestions for improvement?" to receive their qualitative feedback. Once the 7 tasks were completed, participants could easily browse the app at their ease if they wished to discover other functions or pages. To conclude, the experimenter asked a series of sociodemographic questions and questions that included items from two validated scales (System Usability Scale [SUS] [30] and Mobile App Rating Scale [MARS] [31]; Multimedia Appendix 1).

Table 1. List and description of the tasks used during the think-aloud test procedure.

Task name	Expected achievement	Exact wording of the question	Task considered as completed when
Definition	The participant discovers how to find definitions in the app.	“While researching advance directives, you heard the term ‘surrogate’ and want to know what it means exactly. Where would you look in the application to clarify its meaning?”	On the main page of the app, the participant clicks on the term <i>surrogates</i> (highlighted in green because it is a technical term) to pop up the definition; the participant finds the term on the <i>definitions</i> page of the app.
Legal obligations	The participant understands the logic of the green drop-down menu that allows accessing different pages within a section and finding specific legal information.	“You want to know what legal obligation doctors have toward patients who have written advance directives. Where would you go, in this application, to find this information?”	The participant finds the heading “Are doctors required to follow advance directives?” that is included in the page “What does the law say?” situated in the “I get information” section.
Conversation starters	The participant understands that the app is composed of 3 sections accessible via the menu at the bottom of the screen and finds specific content.	“You want to find ways to start a conversation with your son about your advance directives. Where would you go in this application?”	The participant discovered (1) the page <i>It's in the art</i> or (2) the page <i>discussion starters</i> and verbalized that this could be used as a good conversation starter.
Data confidentiality	The participant finds the <i>How it works</i> pop up window that contains key elements for the process of writing advance directives and information on data confidentiality.	“After browsing the application for some time, you feel that you have enough information and decide to start writing your advance directives. However, before you complete the questionnaire, you want to know who will have access to the answers you write in the application. Where would you go in the module to get this information?”	The participant clicks on the button in the <i>I write</i> section and finds the relevant information.
Filter function	The participant discovers the filter function.	“You find that there are too many questions in the questionnaire, and you want to answer in a more targeted way the questions that are important to you. You would like to answer only the questions that concern healthy patients. How would you do this?”	The participant opens the filter function.
Values and preferences	The patient understands that the questionnaire is divided into several subsections and can make sense of their content by reading their title.	“Your faith in God is something central to your life. Therefore, it is really important for you to give details of your religious beliefs. Where would you go in the application to provide details on this matter?”	The participant finds question 15 “A few words about my spirituality” situated in the subsection “My values and preferences” of the questionnaire.
Finalize the questionnaire	The participant understands how to export from the app his or her answers to the questionnaire.	“You have answered all questions that are important to you and now you want to send your form by email. How would you do this?”	The participant clicks on the “Finalize and receive my questionnaire” button situated at the bottom of the questionnaire in the <i>I write</i> section.

Measures Used and Data Analysis

Content Readability

On the basis of a document containing the full content (in text format) of the app, we evaluated its readability, page by page, with the help of Scolarius [32]. Influence Communication, a Canadian media analysis organization, developed this test. The test score is calculated in the same way as the Flesch Reading Ease score (often used for evaluating English material). It calculates the length of words and paragraphs. It provides a score ranging between 50 and 250 to be interpreted as follows: a score between 50 and 89 corresponds to an elementary school

level of education, between 90 and 119 indicates a high school level, between 120 and 149 indicates a college level, and between 150 and 189 indicates a university level.

Tasks Evaluation

Overview

All the video recordings were double coded. CS ensured the first coding of all videos. CC and FE shared the task of the second coding. The recording of one participant was evaluated by 3 authors (CC, CS, and FE) to confirm unanimously that he met an exclusion criterion. Double codes were compared systematically. Whenever coders diverged in their evaluation, we applied the rules listed in Table 2.

Table 2. Rules used to arbitrate diverging evaluations in double coding.

Object of comparison and type of divergence	Rule applied
Task success	
1-point difference between evaluations	Select the lowest score (to avoid desirability bias)
>1 point difference between evaluations	Double check the recording and discuss the evaluation (CS, FE, and CC) until an agreement is reached
Number of clicks	
≥1 point differences between evaluations	CS watches the video again to find the correct number
Time spent on task	
Difference between the 2 evaluations <10%	Use the average time between the 2 evaluators
Difference between the 2 evaluations >10%	CS watches the video again and decides
Type of errors and problems	
Evaluators did not record the same errors and problems encountered by participants	Discuss the evaluation (CS, FE, and CC) until an agreement is reached
Categorization	
Different categorizations for 1 error or problem	Discuss the evaluation (CS, FE, and CC) until an agreement is reached
Severity rate	
Differences between evaluations of the severity rate of an error or problem	Discuss the evaluation (CS, FE, and CC) until an agreement is reached

Task Success

We calculated the number of participants who succeeded or failed to complete the tasks, with or without input from the experimenter. For this, we used the following scoring logic: 0=participant failed because the experimenter eventually gave the answer, 1=succeeded but did not use the shortest path and received some help from the experimenter, 2=succeeded without help but did not use the shortest path, and 3=succeeded without help and easily found the shortest path.

Clicks to Complete the Task

For each task, we reported the number of clicks needed to complete the task. Note that some tasks required more clicks than others; participants did not always start from the same page to complete the task; and for tasks 1 and 3, two answer paths were possible.

Time Spent on Task

For each task, we calculated the number of seconds required by the participants to complete it. Counting started when the experimenter finished formulating the question for the first time and ended when the participant successfully completed the required task (detailed in Table 1, right column). Whenever participants started to digress during task completion (eg, made critical comments on the design of the app or expressed a personal memory), this *digression time* was calculated and subtracted.

Errors and Problems Encountered

We recorded the errors and problems that the participants encountered while completing the 7 tasks. Each problem encountered was described in a Microsoft Excel file, categorized according to the heuristics of Bastien and Scapin [33] (information density, consistency, and significance of codes),

and given a severity rate following the Nielsen Norman Group recommendations [27]: 0=disagreement that this is a usability problem at all, 1=cosmetic problem only—need not be fixed unless extra time is available, 2=minor usability problem: fixing this should be given low priority, 3=major usability problem: important to fix and should be given high priority, and 4=usability catastrophe—imperative to fix this before the product can be released.

Participants' Feedback

Participants' feedback after completing each task was coded as follows: 0=no, the path to complete the task makes no sense to me; 1=yes, it is more or less logical; and 2=yes, it is very logical. We took note of further comments made by the participants during the think-aloud tasks. We only reported comments that (1) were relevant to the objectives of the app and (2) contained proposals or ideas for improvements that could be realistically developed. We categorized comments as *useful for improving the app in the short term* or *useful for future developments*.

Questionnaire

Usability Assessment

To assess the app's usability, we used the 10-item SUS questionnaire [30] suited for evaluating digital products. The SUS contains ten 5-point Likert score questions, allowing the calculation of a total score ranging between 0 and 100 (each question has a 10-point value). This indicates the effectiveness, efficiency, and overall ease of use of this app. A score of approximately 85 is considered excellent, approximately 72 is good, approximately 53 is acceptable, and approximately 38 is poor. Note that the first SUS question ("I think that I would like to use this system frequently") is not relevant for the evaluation of *Accordons-nous*, as discussing end-of-life issues and writing AD is not an everyday activity. To address this issue, we added

a slightly different first question to the 10 original SUS items (1bis), as follows: “I think that if I need to learn about advance care planning and want to write my advance directives, I would give priority to using this application.” This allowed the calculation of two SUS score: one with the standard scale; the other with question 1bis instead of question 1.

Relevance Assessment

To obtain participants’ evaluation of the impact of the app on the user’s awareness, knowledge, attitudes, intentions to change, and the likelihood of actual change in ACP behavior, we used the 6 *perceived impact* items (section F) of the validated MARS questionnaire [31]. The MARS questionnaire is particularly interesting as it was developed on the basis of the transtheoretical model of behavior change by Prochaska [13], which is commonly used in ACP literature to evaluate participants’ change in ACP engagement [34]. It allows calculating the overall mean 5-point Likert scores, to be interpreted as follows: 1=inadequate, 2=poor, 3=acceptable, 4=good, and 5=excellent.

Subjective Endorsement

To observe whether participants endorsed the app, we adapted the first 5-point Likert items of the MARS *subjective quality* (section E) questionnaire. We asked participants whether they would recommend this app to the following three target groups: patients, health professionals, and patients’ caring relatives.

Table 3. Scolarius scores of the app Accordons-nous, per page.

Page of the app	Number of words	Score	Level of education
<i>L’essentiel</i> (key information)	254	100	High school
<i>Histoires de vie</i> (short life stories)	893	94	High school
<i>Que dit la loi?</i> (Swiss legal information)	876	89	Elementary school
<i>Définitions</i> (definitions)	1661	99	High school
<i>Amorces de discussion</i> (discussion starters)	312	60	Elementary school
<i>C’est dans l’art</i> (works of art)	3665	70	Elementary school
<i>I write</i> (filters and questionnaires)	3105	100	High school

Tasks Evaluation

Task Success

In 66% (46/70) of the cases, participants succeeded in completing a task without any help and easily found the shortest path (Figure 4). In 14% (10/70) of the cases, they succeeded without help; however, they did not take the shortest path. In 16% (11/70) of the cases, they succeeded without using the shortest route and needed some help from the experimenter. In 4% (3/70) of the cases, the experimenter eventually gave the answer. This 4% corresponds to 3 cases of failure, involving 30% (3/10) different participants. One failure on the task *definition* was obviously because this participant experienced a temporary state of panic at the beginning of the test and failed

Results

Participants

We obtained responses from 10 participants for all tasks and questions. During the study, we had to exclude 2 more participants: one because of technical problems (the sound recording did not work, making it impossible to analyze the results) and a second as he failed to understand 5 out of the 7 tasks that they were asked to perform, despite 3 iterations of the instructions.

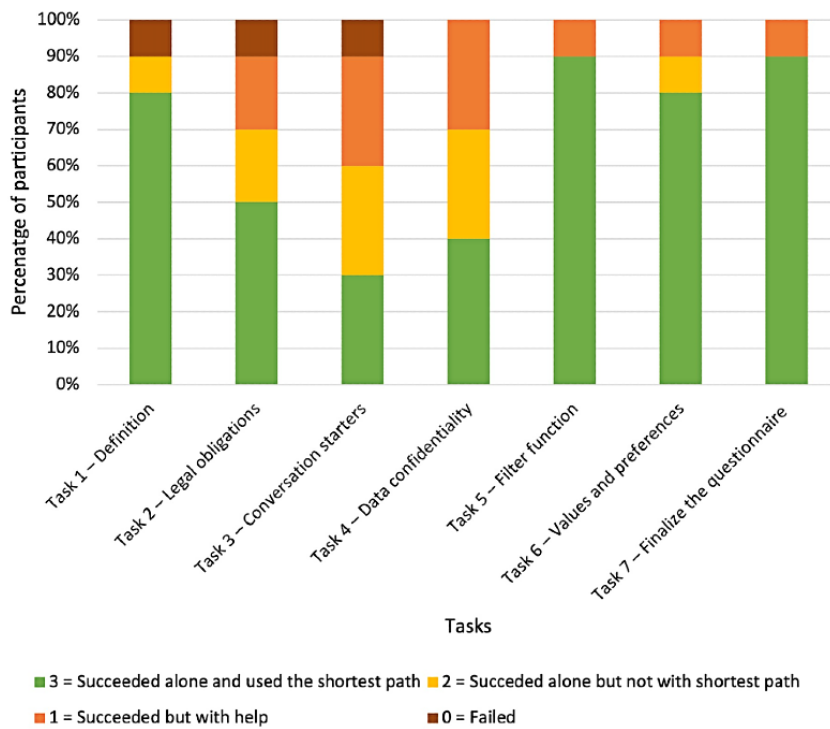
Our data set included 60% (6/10) women and 40% (4/10) men of various ages (youngest: 31 years, oldest: 68 years; mean 50.8, SD 14.6 years). Of the 10 participants, 9 (90%) used devices running on iOS, and 1 (10%) used a device running on Android. All participants used these devices daily and showed minimal understanding of the notion of AD, thereby fulfilling our inclusion criteria.

Content Readability

The overall Scolarius score of the app was 87.43/250, meaning that the app was readable for people with an elementary school level of education. However, some pages of the app require high school-level education. Table 3 provides the details of the score.

to grasp what the experimenter asked them to do (they struggled to find the right definition without looking at the app). In the case of the second failure, on the *legal obligations* task, the participant understood the logic of the drop-down menu within a section and therefore kept switching between the 3 main sections by using the bottom menu. The third failure was on the *conversation starters* task and seemed to be because of participants’ misunderstanding of the meaning of the three main sections: they reported having assumed that sections *I talk about it* and *I write* contained contact forms and therefore thought that this was not the place to find the conversation starters. Furthermore, once they discovered the section *I talk about it* (following guidance from the experimenter), they did not click on the green drop-down menu (a function that they had correctly discovered in task 1).

Figure 4. Proportion of participants who succeeded or failed to complete study tasks with or without input from the experimenter and with or without using the shortest route.

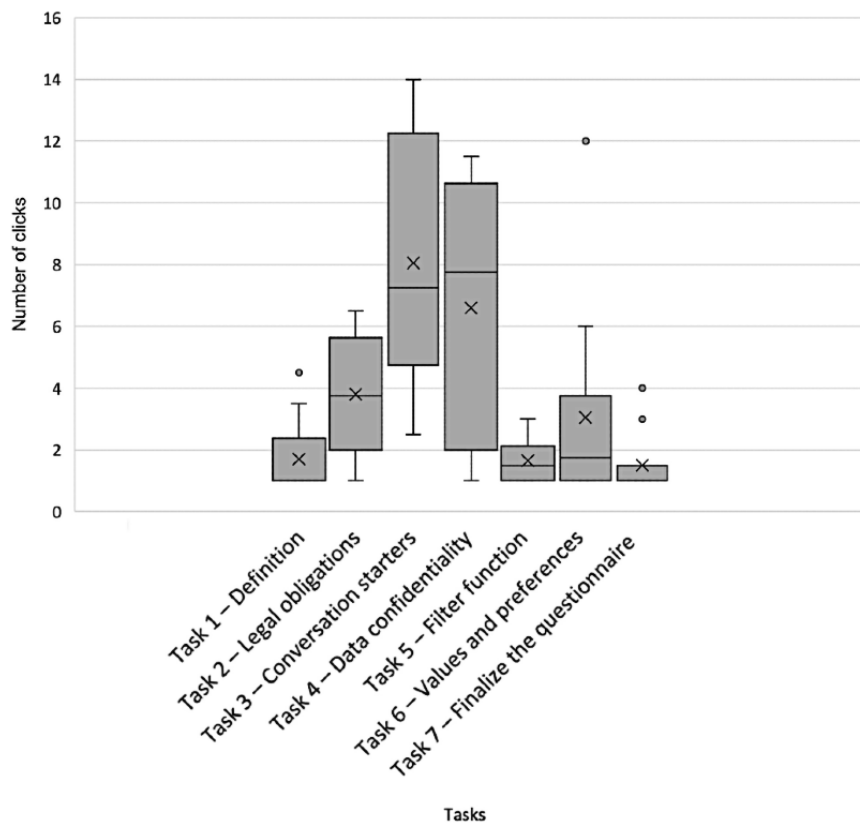


Clicks to Complete the Task

Participants’ actual number of clicks needed is illustrated in Figure 5. During the test, we noticed that the participants struggled to find a direct path toward data confidentiality information (task 4). We wondered whether this difficulty was

affected by the order in which the questions were presented. To check this, in the last 2 interviews of the think-aloud procedure, we changed the order of the questions as follows: 1, 2, 3, 6, 4, 5, and 7. The change in order did not help to complete task 4 faster.

Figure 5. Distribution of the number of clicks needed for participants to complete each of the 7 tasks.



Time Spent on Task

Participants' actual time spent on the tasks is shown in Figure 6. Overall, these results indicate that participants encountered more difficulties in completing the tasks *legal obligations, conversation starters, and data confidentiality*. They encountered little difficulty in completing the last 3 tasks.

Figure 7 shows that there are important differences between participants: some participants navigated without difficulties, whereas others struggled on several tasks. Some participants

needed much more time than others to complete their tasks; for instance, participant 10 used twice more time than participant 7 to successfully complete the tasks. Obviously, participants who needed help from the experimenter took much more time; for instance, participant 3 spent three-fourths of the whole test time on 1 task, for which they needed help.

The participants who succeeded in all tasks (4/10, 40%) were aged 31, 42, 56, and 66 years, indicating that the app is also understandable for older users.

Figure 6. Distribution of time (in seconds) spent by participants completing each of the 7 tasks.

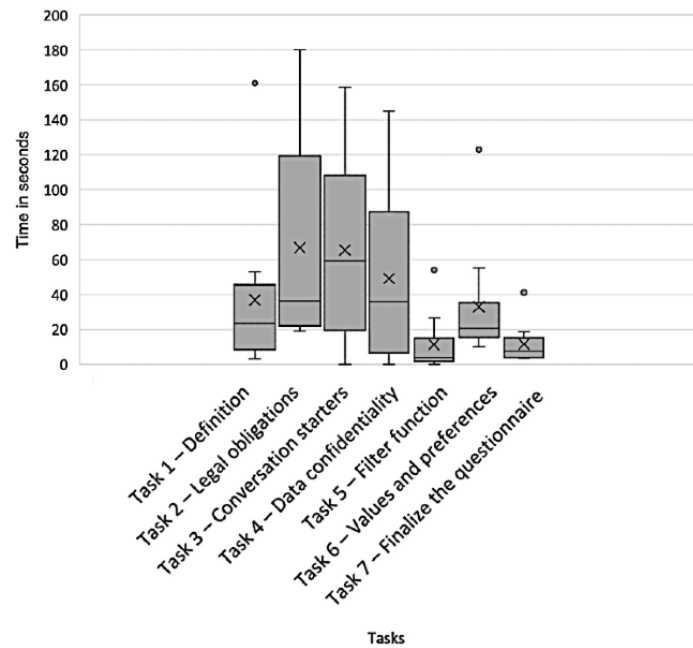
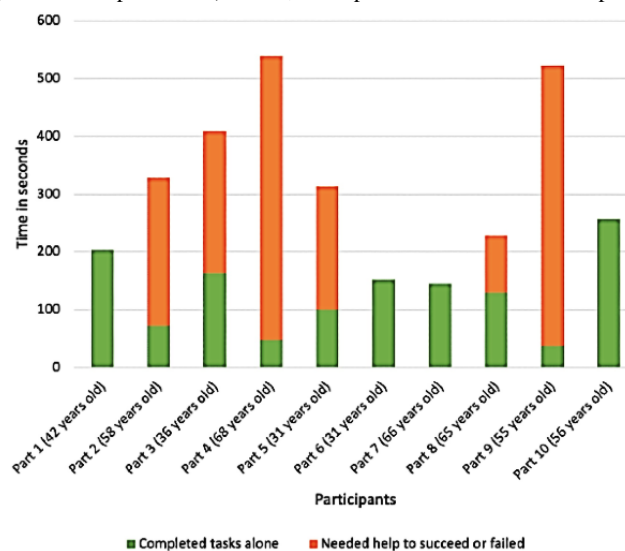


Figure 7. Total time (in seconds) spent by each participant on the 7 tasks of the think-aloud test. In green, time spent on tasks that have been completed successfully and without help (by using the shortest path or not). In red, time spent on tasks failed or completed with the help of the experimenter.



Errors and Problems Encountered

In total, we recorded 44 problems encountered by the participants while completing the 7 tasks. Following the Bastien and Scapin [33] grouping method, we clustered them into 17

types of problems (Table 4). Of these problems, none were considered as a *usability catastrophe*, 50% (22/44) were considered as *major usability problems*, 48% (21/44) as *minor usability problems*, and the last 2% (1/44) as *not a usability problem at all*.

Table 4. List of problems encountered by participants during the 7 think-aloud tasks, classified in 3 severity categories (major usability problem, minor problem, and not a usability issue). Description, categorization, and frequency of occurrence of these problems during task completion.

Item and problem	Category (numbered according to Bastien and Scapin [33])	Description of the problem	Task in which it occurred (frequency of occurrence)
Major usability problems (according to Nielsen Norman Group recommendations [27])			
1	1.2. Grouping or distinction of items	Users did not click on the green drop-down menu and therefore missed much of the content of the module.	Legal obligations (3); Conversation starters (1)
2	7. Significance of codes	Users did not search for the information about the confidentiality of their AD ^a in the right place: they felt that this information should have been located elsewhere in the app (for instance, in <i>I get information</i> , <i>What does the law say</i> , instead of <i>I write</i> , <i>How it works</i> , page).	Data confidentiality (5)
3	1.3. Immediate feedback	Users expected feedback linked to the filter function in the <i>I write</i> section: the module did not provide feedback when a filter was selected. Therefore, users did not know whether their choice was taken into account.	Filter function (4)
4	8. Compatibility	Participants tried to use functions they usually used in other apps; however, these were not active: one iOS user tried swiping left to return to the previous page, and an Android user used the back arrow to achieve the same goal; however, these functions were not active in <i>Concerto</i> .	Conversation starters (1); Data confidentiality (1)
5	2.2 Information density	As the pages contained large sections of written content, some users took time to read the content of the pages, and this delayed their navigation or distracted them from the task. For instance, when they discovered the <i>It's in the art</i> page, some users got <i>stuck</i> to discovering the different works of art.	Legal obligations (1); Conversation starters (2)
6	1.4 Legibility	One user found that the bottom menu is too small and that the font does not offer sufficient contrast.	Conversation starters (1)
7	7. Significance of codes	One user did not find the filter function explicit enough.	Filter function (1)
8	8. Compatibility	One user was confused by the fact that on the first page, the title of the section <i>I get information</i> is larger than the title of the module.	Definition (1)
Minor usability problems			
9	7. Significance of codes	Many users complained that they did not find the expected content by clicking on the hamburger menu at the top right of the screen: they thought that this would allow them to see the architecture of the module. However, this menu is not linked to the module <i>Accordons-nous</i> but to the host app, <i>Concerto</i> .	Definition (1); Legal obligations (2); Conversation starters (3)
10	7. Significance of codes and 8. Compatibility	Many users clicked on the home menu at the top left of the screen, thinking that this would bring them back to the front page of the module <i>Accordons-nous</i> . However, this menu brought the user to the main menu of the host app, <i>Concerto</i> . Users were confused by this response and did not always understand where they had quit <i>Accordons-nous</i> .	Legal obligations (1); Conversation starters (2); Data confidentiality (1)
11	8. Compatibility	Some users were looking for a main <i>home page</i> of the module; however, this page does not exist.	Definition (1); Data confidentiality (1); Values and preferences (1)
12	7. Significance of codes and 8. Compatibility	Two users expected that the <i>I talk about it</i> and <i>I write</i> sections would work as a <i>contact us</i> page, which was not the case.	Conversation starters (2)
13	1.4 Legibility	One user assessed the font size of the text as too small and the font contrast as insufficient.	Definition (1)
14	1.2 Grouping or distinction of items	One user did not see the <i>How does it work?</i> button (section <i>I write</i>).	Data confidentiality (1)
15	7. Significance of codes	One user did not see the <i>Filter function</i> button (section <i>I write</i>).	Filter function (1)
16	8. Compatibility	One ordinary question in the AD form was confused with a function: a user thought that the question, "Where will you save your advance directives?" (multiple options are suggested as an incentive to store AD in several locations) would allow them to save the AD directly in different places.	Finalize questionnaire (1)
Problem that we did not consider to be a usability issue			
17	8. Compatibility	One user complained that the app does not contain a search function.	Definition (1)

^aAD: advance directives.

Participants' Feedback

After completing a task, 64% (45/70) of the time, participants reported that the path to complete the task was *very logical*. In

31% (22/70) of the cases, they reported that it was more or less logical (the remaining 5% did not provide an answer). None of the participants found that the path to complete the task made no sense (Textbox 1).

Textbox 1. List of comments made by participants that are useful in the short and long term for improving the app.

Useful comments for short-term improvements

1. Overall, 40% (4/10) of the oldest participants mentioned that “you have to look everywhere” to understand how to navigate.
2. Data privacy was a recurrent concern.
3. In total, 30% (3/10) of participants pointed out that the font was very small, which may be problematic for elderly users.
4. Many participants found that there might be too much to read in the app, especially given the fact that people tend to skip lengthy paragraphs.
5. Many participants said that they liked the example responses provided in the questionnaire (for illustration, see Figure 8). However, to avoid overloading the module with text, we included these texts in a lighter gray within the response box. The text disappears once the user clicks on the text box to add his or her own response. Some participants would have preferred these examples to remain visible (be placed above the box) to “not forget anything.”

Useful comments for long-term improvements

1. Some participants would have appreciated a search function.
2. With this app, a participant expected to be able to make calls or contact people around him or health professionals at the Hopitaux Universitaires de Genève.
3. A participant wished to have the function of electronic signature for signing the advance directives on the app.
4. A participant wished for more options for disabled patients, such as a dictation function for people who cannot write.

Figure 8. Two questions included in the advance directives (AD) form, with example responses in gray. The text disappears once the user clicks on the text box to add his or her own response.

11 - LA OU LES CHOSE(S) QUE J'AIMERAIS ABSOLUMENT FAIRE AVANT DE MOURIR

Ex: découvrir un lieu, faire un voyage, une fête, concrétiser un souhait, avoir une conversation, clore une affaire non réglée, me réconcilier avec quelqu'un, etc.

12 - CE QUI ME FAIT PEUR, ME PRÉOCUPE, M'ANGOISSE

Ex: que vont devenir les personnes qui dépendent de moi? Vont-elles souffrir? Est-ce que je risque de devenir une charge pour ma famille? Que vont devenir mes animaux de compagnie? Mes affaires financières sont-elles en ordre? Est-il possible qu'on ne me dise pas tout? Ai-je encore la maîtrise de ma vie? Vais-je souffrir? etc

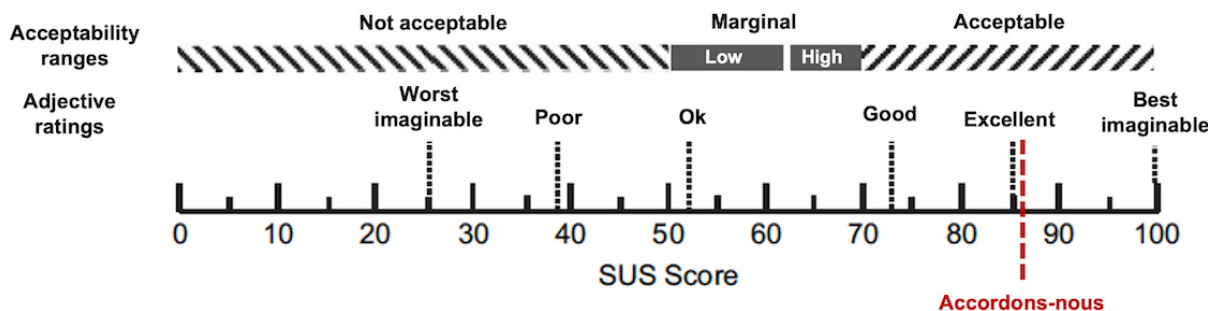
Questionnaire

Usability Score

As illustrated in Figure 9 [35], the app *Accordons-nous* scores very high on the SUS usability scale (85.25/100). When question

1 of the SUS was replaced with question 1bis, the total score increased to 90.5.

Figure 9. Score of Accordons-nous on the SUS. SUS: System Usability Scale [35].



Relevance Score

Participants’ evaluation of the relevance of the app for supporting the process of ACP and AD is reported in Table 5. Most notably, 90% (9/10) to 100% (10/10) of the participants

agree or strongly agree that *Accordons-nous* is likely to increase knowledge about the topic and the motivation to address the topic and induce related behavior changes. These responses lead to a high adherence rate, with an overall score of 4.27/5.

Table 5. Participants’ evaluation of the relevance of Accordons-nous for improving ACP^a and AD^b (N=10).

Measure	Question item	Responses, n (%)				
		Strongly agree	Agree	No opinion	Disagree	Strongly disagree
Awareness	This app is likely to increase awareness of the importance of (ACP and AD)	5 (50)	4 (40)	0 (0)	1 (10)	0 (0)
Knowledge	This app is likely to increase knowledge or understanding of (ACP and AD)	9 (90)	1 (10)	0 (0)	0 (0)	0 (0)
Attitude	This app is likely to change attitudes toward improving (ACP and AD)	3 (30)	4 (40)	3 (30)	0 (0)	0 (0)
Intention to change	This app is likely to increase intentions or motivation to address (ACP and AD)	4 (40)	5 (50)	1 (10)	0 (0)	0 (0)
Help seeking	Using this app is likely to encourage further help seeking for (ACP and AD)	1 (10)	5 (50)	3 (30)	1 (10)	0 (0)
Behavior change	Using this app is likely to increase behavior change (ACP and AD)	4 (40)	6 (60)	0 (0)	0 (0)	0 (0)

^aACP: advance care planning.

^bAD: advance directives.

Subjective Endorsement

Overall, participants endorsed *Accordons-nous*. Of the 10 participants, 9 (90%) recommended the app (2 times 4/5 and 7 times 5/5 on the Likert scale) for patients. Only 10% (1/10) of participants *strongly disagreed* (1/5, lowest score) to recommend the app to patients, arguing that a smartphone app is not suitable for an older audience, whom they found as not being at ease with such technology. However, 90% (9/10) participants *strongly recommended* (5/5) the app to health professionals. Only 10% (1/10) disagreed (2/5), arguing that the tool was primarily designed for patients, and therefore, a professional

would not find it relevant to their practice. Finally, all participants recommended (1 time 4/5 and 9 times 5/5) the app to family caregivers. The mean endorsement over the 3 questions was 4.7/5.

Modifications Made Based on Results

The usability test revealed a series of issues that were addressed, as described in Table 6. *Accordons-nous* is an app in an app; that is, a module inserted in the larger host app *Concerto*. For this reason, some navigability issues could only be addressed by the *Concerto* team. Table 7 describes the issues that we reported to *Concerto* and how they have been addressed.

Table 6. List of changes made to the app in response to the major issues identified.

Item (Table 4)	Description of the issue	Changes made
1 and 6	Users did not click on the green drop-down menu and therefore missed much of the content of the module. One user found that the bottom menu is too small and that the font does not offer sufficient contrast.	To overcome navigation difficulties, we created a motion design video [36] (inserted on the first page of the app) containing navigation tutorial elements.
2	Users did not search for the information about the confidentiality of their advance directives in the right place: they felt that this information should have been located elsewhere in the app (for instance, in the <i>I get information</i> , <i>What does the law say</i> , instead of the <i>I write</i> , <i>How it works</i> , page).	We duplicated the information on “How is your privacy managed?” on the first page of the module.
3 and 7	Users expected feedback linked to the filter function in the <i>I write</i> section: the module did not provide feedback when a filter was selected; therefore, users did not know whether their choice was taken into account. One user did not find the filter function explicit enough.	First, we changed the color and location of the <i>How it works</i> button to make it more visible and added a video tutorial to explain how to write and store advance directives [23]. Second, we redesigned the filter page: we renamed and relocated the drop-down filter button to make it more explicit that it is a filter menu. Third, we added an automatic notification after the user has made a choice of filter.
5	As pages contained large sections of written content, some users took time to read the content of the pages, and this delayed their navigation or distracted them from the task. For instance, when they discovered the <i>It's in the art</i> page, some users got <i>stuck</i> on discovering the different works of art.	First, we integrated 2 motion design videos containing illustrated summaries of the most important information. Second, we used the accordion visual presentation in sections containing large amounts of texts to help users to find the searched information without having to scroll through long texts. Third, in the <i>I talk about it</i> section, we preselected the entry page that contains less text to avoid a feeling of overload while navigating across sections.
8	One user was confused by the fact that on the first page, the title of the section <i>I get information</i> is larger than the title of the module.	We removed the title of the app on all pages.

Table 7. List of problems that were notified to the Concerto team and changes made.

Item (Table 4)	Reported problem	Change made
9	Several users clicked on the hamburger menu at the top right of the screen, thinking that this would allow them to see the architecture of <i>Accordons-nous</i> . However, this button is linked to the host app <i>Concerto</i> and provides only log-in information, which is confusing for the user.	The <i>Concerto</i> team removed the hamburger menu in <i>Accordons-nous</i> except when users are logged
10	Several users clicked on the home button at the top left of the screen, thinking that this would allow them to go back to the front page of <i>Accordons-nous</i> . However, this button brings the user to the front page of <i>Concerto</i> . This action involved quitting <i>Accordons-nous</i> , which also confused some of the users.	The <i>Concerto</i> team changed the design of the button: instead of a house, it is now the pictogram of the <i>Concerto</i> home page
4	Participants tried to use functions they usually used in other apps; however, these were not active: an iOS user tried swiping left to return to the previous page, and an Android user used the back arrow to achieve the same goal; however, these functions were not active in <i>Concerto</i> .	None to this date
13	The <i>Concerto</i> text font (imposed to all modules, including <i>Accordons-nous</i>) was judged as too small and not dark enough (most text is in a shade of gray).	None to this date

Discussion

Principal Findings

Accordons-nous is the first French-language mobile app for supporting the ACP process. In contrast to existing apps worldwide, which are limited in many respects [21], it includes materials and functions that facilitate each stage of the ACP process. It provides all the relevant information in an accessible language. It includes tools and prompts to facilitate contemplation and discussions about the goals of care and end-of-life issues. It provides guidance on writing or updating comprehensive and personalized AD in a simple process. In

line with professional recommendations [22], it can be used as an icebreaker for starting discussions within families and as a follow-up to direct conversations with health professionals.

In our usability test, participants succeeded in the 7 think-aloud tasks in 80% (56/70) of the cases without help from the experimenter. However, in 16% (11/70) of the cases, help was needed, indicating that the app needed some improvement. An analysis of the navigational errors and difficulties encountered revealed no usability catastrophe but a series of major usability problems. These can be addressed with minor modifications (Table 6).

Our graphs indicate that participants performed better at completing the tasks over the course of the test, although the latter tasks were not necessarily easier to complete. We think that participants gradually became familiar and at ease with the logic of the app, which indicates that a few minutes of use are enough to grasp our tool. The easy-to-use character of the tool is particularly important [15] as ACP is mainly relevant for older patients [3] who are less used to digital devices. Our oldest participants (aged 66 and 68 years) fulfilled the think-aloud tasks well, indicating that age was not necessarily a limiting factor for using our tool. This is in line with previous results indicating that older patients are likely to use a platform for ADs with adequate design and support [15]. Moreover, as we designed *Accordons-nous* as a tool to help engage in discussions with family and health care providers, even patients who are not accustomed to digital devices may obtain support from surrounding care providers.

The app *Accordons-nous* obtained high scores on readability (overall score of 87.4 on the Scolarius test, corresponding to elementary school level). This is one of the best scores compared with other similar apps internationally [21]. However, the language difficulty varies across different pages of the app. Owing to the inherent complexity of the ACP topic, this difficulty can only be partially addressed; we ensured that all technical terms in the app were clickable to pop up the definition and added 2 introductory motion design videos summarizing the main information in simple terms and with illustrations.

Overall, our users provided very positive feedback on *Accordons-nous*. They assessed the app as likely to raise awareness of the importance of ACP and AD, increase knowledge of such topics, change attitudes and behaviors intention toward improving their own ACP, and encourage seeking further help to fulfill one's ACP. Moreover, participants expressed a clear willingness to recommend the tool to relevant stakeholders, including patients, health professionals, and patients' caring relatives. These are exciting results indicating that *Accordons-nous* is relevant and contains easily understandable content.

However, among the few critical views expressed, it is worth noting that 10% (1/10) of patients disagreed that *Accordons-nous* may encourage further help seeking in engaging in ACP, and 30% (3/10) had no opinion on that question. This result may indicate that patients expect the tool to offer the possibility of sending direct calls for help (eg, sending a message to a task force of professional ACP facilitators). Indeed, during the test, 20% (2/10) of participants erroneously thought at first glance

that the section *I talk about it* would provide such direct support service. For practical and organizational reasons, we were not able to provide direct access to private counseling with *Accordons-nous*, which is a limitation of the app.

Our analysis of the main problems encountered by participants while completing the 7 think-aloud tasks indicates that some participants struggled with cognitive load: they failed a task or needed help to complete it as their attention was caught by the high density of information contained in some pages of the app. For instance, some participants got distracted by the long list of works of art that they discovered on the page *works of art*.

Other difficulties were because some participants did not spontaneously grasp the logic of navigating the app. In particular, some participants did not immediately see that the app contained 3 sections accessible via the navigation bar at the bottom of the screen (this was needed to complete the task *conversation starters*), whereas other participants encountered difficulties with the logic of the drop-down menu that allows navigating different pages (this was needed to complete tasks *legal obligation* and *conversation starters*). These difficulties indicate that some improvements needed to be made to the tool, especially for users who are not familiar with smartphone apps.

To address the main navigation difficulties and cognitive load because of the high information content, we included 2 introductory motion design videos in the app. In the first video [36], we explained the aim of the app, why ACP is useful, how the app is structured (3 sections), and the main navigation functions (menu to navigate the 3 main sections and drop-down menu buttons). In the second video [23], we explained how confidentiality issues are managed and how to write and store ADs with or without the help of the app.

Regarding task *data confidentiality*, several participants struggled to find the information as they searched it in the *I get informed* section rather than in the *I write* section. For this task, half of the participants reported that the path to complete it was *more or less logical* (rather than *very logical*). This indicates that data about confidentiality should be duplicated in both sections to be more easily found. This is particularly important as previous studies indicate that users are concerned about privacy issues [37]. We adapted the content of the app accordingly (Table 6).

Our results must be interpreted while considering the limitations listed in Textbox 2. Some limitations are common to usability studies, whereas others are minor issues related to small deviations from the original study protocol.

Textbox 2. List of identified limitations to our study.

Choice of method

- As our test includes a small number of participants, no statistical conclusion can be drawn from this study.

Sampling method

- Owing to our sampling method (a small convenience selection of participants from the *Patients as Partners Project* at Hopitaux Universitaires de Genève), a courtesy bias [38] may have inflated positive results. Therefore, it is important to pay particular attention to the critical feedback provided.
- It may be that participants were more trained in medical matters than ordinary users, and this may have affected our results. On the other hand, ordinary patients would have much more time to browse and discover the app than our participants.

Test conditions

- Some failures or difficulties that we recorded may be because of the lack of time and stress induced by the test.

Deviation from the study protocol

- While writing the original protocol, we did not specify the exclusion criterion “Incapable of understanding the meaning of the questions and tasks required for the study.” However, we decided to exclude one participant on this criterion as he obviously did not grasp most of the tasks that we asked him to complete.
- During the test, we changed the order of the questions from 1-2-3-4-5-6-7 to 1-2-3-6-4-5-7 (see explanation in section *Results*, subsection *Clicks to Complete the Task*) The change of order did not help to complete task 4 faster. This change of order may have influenced the number of clicks to complete the tasks.

Conclusions

The tool we developed is a novel solution for promoting ACP and AD. *Accordons-nous* is the first French-language mobile app developed by an interdisciplinary team of professionals in collaboration with target users. It includes a variety of content for prompting discussions related to medical emergency situations and end-of-life issues. It provides support for writing and easily updating AD on a smartphone or tablet. Considering

the complexity and sensitivity of the process of ACP and given that we expect most users to be older people, we put special emphasis on producing easy-to-understand information, discussion prompts, and simple navigation principles. The results of our usability test with patients were very satisfying and helped us make the necessary final adjustments to our tool before making it available to the public. Further usability and efficacy tests involving health care professionals would help define whether the tool is also suitable for this population.

Acknowledgments

The authors would like to thank everyone who contributed to this project. In particular, The authors thank Inès Serre and Pierre Sutter (members of the Hopitaux Universitaires de Genève [HUG] *Patients as Partners Project*) for their availability and helpful critical feedback throughout the project; François Pierre Robert and Vincent Clavien (information technology professionals) for helping us to develop the first prototype; Antoine Geissbühler and Helena Bornet Dit Vorgeat (HUG Innovation Center), Marianne Rubio and Simon Schmitt (HUG Communication Department), André Koller and Catherine Chenaud (HUG patient and family information group) for their regular support and counseling throughout the project; the many professionals who provided critical feedback at different stages of content development, notably Samia Hurst, Bara Ricou, Thierry Fumeaux, Louise Penzenstadler, Ralf Jox, Sophie Pautex, Monika Escher, Marinette Ummel, Francesca Bosisio, Johanna Sommer, Katherine Blondon, Véronique Dauve, Gora Da Rocha, Sophie Courvoisier, Laurence Sechaud, Tanja Krones, Julia Sader, Beat Sottas, Catherine Hoenerger, Alex Mauron; and our 11 anonymous study participants.

Conflicts of Interest

Our project benefited from the financial support of the Fondation Privée des Hopitaux Universitaires de Genève [39], which is a nonprofit organization that was not involved in any step of this study. *Accordons-nous* is not a saleable product. It is included in the host app Concerto Hopitaux Universitaires de Genève, which can be downloaded for free. The authors have no financial interests to declare.

Multimedia Appendix 1

The history of the development of *Accordons-nous* and the list of questions submitted to the participants.

[[DOCX File , 44 KB - humanfactors_v9i2e34626_app1.docx](#)]

References

1. McMahan RD, Tellez I, Sudore RL. Deconstructing the complexities of advance care planning outcomes: what do we know and where do we go? A scoping review. *J Am Geriatr Soc* 2021 Jan 07;69(1):234-244 [FREE Full text] [doi: [10.1111/jgs.16801](https://doi.org/10.1111/jgs.16801)] [Medline: [32894787](https://pubmed.ncbi.nlm.nih.gov/32894787/)]
2. La planification anticipée concernant la santé, en particulier en cas d'incapacité de discernement (« Advance Care Planning »). Cadre général pour la Suisse. Département fédéral de l'intérieur DFI. URL: https://www.plateforme-soinspalliatifs.ch/sites/default/files/work/files/Rahmenkonzept_Gesundheitl_Vorausplanung_FR_rev.pdf [accessed 2022-03-24]
3. Rietjens JA, Sudore RL, Connolly M, van Delden JJ, Drickamer MA, Droger M, et al. Definition and recommendations for advance care planning: an international consensus supported by the European Association for Palliative Care. *Lancet Oncol* 2017 Sep;18(9):e543-e551. [doi: [10.1016/S1470-2045\(17\)30582-X](https://doi.org/10.1016/S1470-2045(17)30582-X)]
4. Walsh E. Cognitive transformation, dementia, and the moral weight of advance directives. *Am J Bioeth* 2020 Aug 06;20(8):54-64. [doi: [10.1080/15265161.2020.1781955](https://doi.org/10.1080/15265161.2020.1781955)] [Medline: [32757910](https://pubmed.ncbi.nlm.nih.gov/32757910/)]
5. Bosisio F, Jox RJ, Jones L, Rubli Truchard E. Planning ahead with dementia: what role can advance care planning play? A review on opportunities and challenges. *Swiss Med Wkly* 2018 Dec 17;148:w14706 [FREE Full text] [doi: [10.4414/smw.2018.14706](https://doi.org/10.4414/smw.2018.14706)] [Medline: [30594990](https://pubmed.ncbi.nlm.nih.gov/30594990/)]
6. Vilpert S, Borrat-Besson C, Maurer J, Borasio GD. Awareness, approval and completion of advance directives in older adults in Switzerland. *Swiss Med Wkly* 2018 Jul 16;148:w14642 [FREE Full text] [doi: [10.4414/smw.2018.14642](https://doi.org/10.4414/smw.2018.14642)] [Medline: [30141525](https://pubmed.ncbi.nlm.nih.gov/30141525/)]
7. Cattagni Kleiner A, Santos-Eggimann B, Fustinoni S, Dürst AV, Haunreiter K, Rubli-Truchard E, et al. Advance care planning dispositions: the relationship between knowledge and perception. *BMC Geriatr* 2019 Apr 24;19(1):118 [FREE Full text] [doi: [10.1186/s12877-019-1113-3](https://doi.org/10.1186/s12877-019-1113-3)] [Medline: [31014271](https://pubmed.ncbi.nlm.nih.gov/31014271/)]
8. Schickedanz A, Schillinger D, Landefeld C, Knight S, Williams B, Sudore R. A clinical framework for improving the advance care planning process: start with patients' self-identified barriers. *J Am Geriatr Soc* 2009 Jan;57(1):31-39 [FREE Full text] [doi: [10.1111/j.1532-5415.2008.02093.x](https://doi.org/10.1111/j.1532-5415.2008.02093.x)] [Medline: [19170789](https://pubmed.ncbi.nlm.nih.gov/19170789/)]
9. Pautex S, Herrmann F, Zulian G. Role of advance directives in palliative care units: a prospective study. *Palliat Med* 2008 Oct 28;22(7):835-841. [doi: [10.1177/0269216308094336](https://doi.org/10.1177/0269216308094336)] [Medline: [18718993](https://pubmed.ncbi.nlm.nih.gov/18718993/)]
10. Periyakoil V, Neri E, Kraemer H. No easy task: a mixed methods study of barriers to conducting effective end-of-life conversations reported by multi-specialty doctors (S776). *J Pain Symptom Manag* 2015 Feb;49(2):446-447. [doi: [10.1016/j.jpainsymman.2014.11.255](https://doi.org/10.1016/j.jpainsymman.2014.11.255)]
11. Morrison R, Morrison E, Glickman D. Physician reluctance to discuss advance directives. An empiric investigation of potential barriers. *Arch Intern Med* 1994 Oct 24;154(20):2311-2318. [Medline: [7944853](https://pubmed.ncbi.nlm.nih.gov/7944853/)]
12. Fried TR, Redding CA, Robbins ML, Paiva A, O'Leary JR, Iannone L. Promoting advance care planning as health behavior change: development of scales to assess Decisional Balance, Medical and Religious Beliefs, and Processes of Change. *Patient Educ Couns* 2012 Jan;86(1):25-32 [FREE Full text] [doi: [10.1016/j.pec.2011.04.035](https://doi.org/10.1016/j.pec.2011.04.035)] [Medline: [21741194](https://pubmed.ncbi.nlm.nih.gov/21741194/)]
13. Prochaska JO, Velicer WF. The transtheoretical model of health behavior change. *Am J Health Promot* 1997 Sep;12(1):38-48. [doi: [10.4278/0890-1171-12.1.38](https://doi.org/10.4278/0890-1171-12.1.38)] [Medline: [10170434](https://pubmed.ncbi.nlm.nih.gov/10170434/)]
14. Woolfolk RL, Doris JM, Darley JM. Identification, situational constraint, and social cognition: studies in the attribution of moral responsibility. *Cognition* 2006 Jun;100(2):283-301. [doi: [10.1016/j.cognition.2005.05.002](https://doi.org/10.1016/j.cognition.2005.05.002)] [Medline: [16087171](https://pubmed.ncbi.nlm.nih.gov/16087171/)]
15. Portz JD, Lum HD, Bull S, Boxer RS, Bekelman DB, Ford KL, et al. Perceptions of patient portal use for advance directive documentation among older adults with multiple chronic conditions. *J Soc Work End Life Palliat Care* 2020 Jun 01;16(3):238-249. [doi: [10.1080/15524256.2020.1771806](https://doi.org/10.1080/15524256.2020.1771806)] [Medline: [32482160](https://pubmed.ncbi.nlm.nih.gov/32482160/)]
16. Fine RL, Yang Z, Spivey C, Boardman B, Courtney M. Early experience with digital advance care planning and directives, a novel consumer-driven program. *Proc (Bayl Univ Med Cent)* 2016 Jul 11;29(3):263-267 [FREE Full text] [doi: [10.1080/08998280.2016.11929430](https://doi.org/10.1080/08998280.2016.11929430)] [Medline: [27365867](https://pubmed.ncbi.nlm.nih.gov/27365867/)]
17. van der Smissen D, Overbeek A, van Dulmen S, van Gemert-Pijnen L, van der Heide A, Rietjens JA, et al. The feasibility and effectiveness of web-based advance care planning programs: scoping review. *J Med Internet Res* 2020 Mar 17;22(3):e15578 [FREE Full text] [doi: [10.2196/15578](https://doi.org/10.2196/15578)] [Medline: [32181750](https://pubmed.ncbi.nlm.nih.gov/32181750/)]
18. Sudore RL, Boscardin J, Feuz MA, McMahan RD, Katen MT, Barnes DE. Effect of the PREPARE website vs an easy-to-read advance directive on advance care planning documentation and engagement among veterans: a randomized clinical trial. *JAMA Intern Med* 2017 Aug 01;177(8):1102-1109 [FREE Full text] [doi: [10.1001/jamainternmed.2017.1607](https://doi.org/10.1001/jamainternmed.2017.1607)] [Medline: [28520838](https://pubmed.ncbi.nlm.nih.gov/28520838/)]
19. Howard M, Robinson CA, McKenzie M, Fyles G, Sudore RL, Andersen E, et al. Effect of an interactive website to engage patients in advance care planning in outpatient settings. *Ann Fam Med* 2020 Mar 09;18(2):110-117 [FREE Full text] [doi: [10.1370/afm.2471](https://doi.org/10.1370/afm.2471)] [Medline: [32152014](https://pubmed.ncbi.nlm.nih.gov/32152014/)]
20. Sudore RL, Knight SJ, McMahan RD, Feuz M, Farrell D, Miao Y, et al. A novel website to prepare diverse older adults for decision making and advance care planning: a pilot study. *J Pain Symptom Manag* 2014 Apr;47(4):674-686 [FREE Full text] [doi: [10.1016/j.jpainsymman.2013.05.023](https://doi.org/10.1016/j.jpainsymman.2013.05.023)] [Medline: [23972574](https://pubmed.ncbi.nlm.nih.gov/23972574/)]
21. McDarby M, Llaneza D, George L, Kozlov E. Mobile applications for advance care planning: a comprehensive review of features, quality, content, and readability. *Am J Hosp Palliat Care* 2021 Aug 17;38(8):983-994 [FREE Full text] [doi: [10.1177/1049909120959057](https://doi.org/10.1177/1049909120959057)] [Medline: [32938207](https://pubmed.ncbi.nlm.nih.gov/32938207/)]

22. MacKenzie Greenle M, Morgan B, Sayani S, Meghani SH. Identifying mobile apps targeting palliative care patients and family members. *J Palliat Med* 2018 Oct;21(10):1380-1385. [doi: [10.1089/jpm.2018.0157](https://doi.org/10.1089/jpm.2018.0157)] [Medline: [30312126](https://pubmed.ncbi.nlm.nih.gov/30312126/)]
23. HUG Anticipation 02. YouTube. URL: <https://www.youtube.com/watch?v=j4FF56f0XuY> [accessed 2022-03-24]
24. Hackathons des HUG. HUG - Hôpitaux Universitaires Genève. URL: <https://www.hug.ch/centre-innovation/hackathon> [accessed 2022-04-05]
25. What is Patients as Partners? HUG - Hôpitaux Universitaires Genève. URL: <https://www.hug.ch/en/patients-partenaires/what-patients-partners> [accessed 2022-04-05]
26. Jain N, Mangal P, Mehta D. AngularJS: a modern MVC framework in JavaScript. *J Global Res Comput Sci* 2014;5(2):17-23 [FREE Full text]
27. Nielson J, Landauer TK. A mathematical model of the finding of usability problems. In: Proceedings of the INTERACT '93 and CHI '93 Conference on Human Factors in Computing Systems. 1993 Presented at: INTERCHI93: Conference on Human Factors in Computing; Apr 24 - 29, 1993; Amsterdam, The Netherlands. [doi: [10.1145/169059.169166](https://doi.org/10.1145/169059.169166)]
28. Albert W, Tullis T, Albert W. Measuring the User Experience Collecting, Analyzing, and Presenting Usability Metrics. Amsterdam, Netherlands: Elsevier Science; 2010.
29. Jaspers MW, Steen T, van den Bos C, Geenen M. The think aloud method: a guide to user interface design. *Int J Med Inform* 2004 Nov;73(11-12):781-795. [doi: [10.1016/j.ijmedinf.2004.08.003](https://doi.org/10.1016/j.ijmedinf.2004.08.003)] [Medline: [15491929](https://pubmed.ncbi.nlm.nih.gov/15491929/)]
30. Brooke J. SUS: a 'quick and dirty' usability scale. In: Usability Evaluation In Industry. Boca Raton, Florida: CRC Press; 1996.
31. Stoyanov SR, Hides L, Kavanagh DJ, Zelenko O, Tjondronegoro D, Mani M. Mobile app rating scale: a new tool for assessing the quality of health mobile apps. *JMIR Mhealth Uhealth* 2015 Mar 11;3(1):e27 [FREE Full text] [doi: [10.2196/mhealth.3422](https://doi.org/10.2196/mhealth.3422)] [Medline: [25760773](https://pubmed.ncbi.nlm.nih.gov/25760773/)]
32. Scolaris homepage. Scolarius. URL: <https://www.scolarius.com/> [accessed 2022-03-24]
33. Bastien JC, Scapin DL. A validation of ergonomic criteria for the evaluation of human - computer interfaces. *Int J Human Comput Interact* 1992 Apr;4(2):183-196. [doi: [10.1080/10447319209526035](https://doi.org/10.1080/10447319209526035)]
34. Van Scoy LJ, Green MJ, Reading JM, Scott AM, Chuang CH, Levi BH. Can playing an end-of-life conversation game motivate people to engage in advance care planning? *Am J Hosp Palliat Care* 2017 Sep 12;34(8):754-761 [FREE Full text] [doi: [10.1177/1049909116656353](https://doi.org/10.1177/1049909116656353)] [Medline: [27406696](https://pubmed.ncbi.nlm.nih.gov/27406696/)]
35. Bangor A, Kortum P, Miller J. Determining what individual SUS scores mean: adding an adjective rating scale. *J Usab Stud* 2009;4(3):114-123. [doi: [10.5555/2835587.2835589](https://doi.org/10.5555/2835587.2835589)]
36. HUG Anticipation 01. YouTube. URL: <https://www.youtube.com/watch?v=KTKYsOPyNMQ> [accessed 2022-03-24]
37. Sakaguchi-Tang DK, Bosold AL, Choi YK, Turner AM. Patient portal use and experience among older adults: systematic review. *JMIR Med Inform* 2017 Oct 16;5(4):e38 [FREE Full text] [doi: [10.2196/medinform.8092](https://doi.org/10.2196/medinform.8092)] [Medline: [29038093](https://pubmed.ncbi.nlm.nih.gov/29038093/)]
38. Evan WM, Miller JR. Differential effects on response bias of computer vs. conventional administration of a social science questionnaire: an exploratory methodological experiment. *Syst Res* 1969 May;14(3):216-227. [doi: [10.1002/bs.3830140306](https://doi.org/10.1002/bs.3830140306)]
39. Fondation privée des HUG. URL: <https://www.fondationhug.org/> [accessed 2022-03-24]

Abbreviations

- ACP:** advance care planning
AD: advance directives
HUG: Hopitaux Universitaires de Genève
MARS: Mobile App Rating Scale
SUS: System Usability Scale

Edited by A Kushniruk; submitted 01.11.21; peer-reviewed by C Dupont, R Marshall; comments to author 05.12.21; revised version received 11.01.22; accepted 20.01.22; published 20.04.22.

Please cite as:

Schöpfer C, Ehrler F, Berger A, Bollondi Pauly C, Buytaert L, De La Serna C, Hartheiser F, Fassier T, Clavien C
A Mobile App for Advance Care Planning and Advance Directives (Accordons-nous): Development and Usability Study
JMIR Hum Factors 2022;9(2):e34626
URL: <https://humanfactors.jmir.org/2022/2/e34626>
doi: [10.2196/34626](https://doi.org/10.2196/34626)
PMID: [35442206](https://pubmed.ncbi.nlm.nih.gov/35442206/)

20.04.2022. This is an open-access article distributed under the terms of the Creative Commons Attribution License (<https://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Human Factors, is properly cited. The complete bibliographic information, a link to the original publication on <https://humanfactors.jmir.org>, as well as this copyright and license information must be included.

Original Paper

Understanding the Potential of Mental Health Apps to Address Mental Health Needs of the Deaf and Hard of Hearing Community: Mixed Methods Study

Judith Borghouts¹, PhD; Martha Neary², MSc; Kristina Palomares², BA; Cinthia De Leon¹, MPH; Stephen M Schueller^{2,3}, PhD; Margaret Schneider⁴, PhD; Nicole Stadnick^{5,6,7}, MPH, PhD; Dana B Mukamel¹, PhD; Dara H Sorkin¹, PhD; Dakota Brown⁸; Shannon McCleerey-Hooper⁸, BA; Gloria Moriarty⁹; Elizabeth V Eikey^{10,11}, PhD

¹Department of Medicine, University of California, Irvine, Irvine, CA, United States

²Department of Psychological Science, University of California, Irvine, Irvine, CA, United States

³Department of Informatics, University of California, Irvine, Irvine, CA, United States

⁴Department of Public Health, University of California, Irvine, Irvine, CA, United States

⁵Department of Psychiatry, University of California, San Diego, La Jolla, CA, United States

⁶Dissemination and Implementation Science Center, Altman Clinical and Translational Research Institute, University of California, San Diego, La Jolla, CA, United States

⁷Child and Adolescent Services Research Center, San Diego, CA, United States

⁸Riverside University Health System-Behavioral Health, Riverside, CA, United States

⁹Center on Deafness Inland Empire, Riverside, CA, United States

¹⁰Herbert Wertheim School of Public Health and Human Longevity Science, University of California, San Diego, La Jolla, CA, United States

¹¹The Design Lab, University of California, San Diego, La Jolla, CA, United States

Corresponding Author:

Judith Borghouts, PhD
Department of Medicine
University of California, Irvine
100 Theory
Irvine, CA, 92617
United States
Phone: 1 9498240246
Email: jborghou@uci.edu

Abstract

Background: Mental health concerns are a significant issue among the deaf and hard of hearing (D/HH) community, but community members can face several unique challenges to accessing appropriate resources.

Objective: The aim of this study was to investigate the mental health needs of the D/HH community and how mental health apps may be able to support these needs.

Methods: A total of 10 members of the D/HH community participated in a focus group and survey to provide their perspectives and experiences. Participants were members of the Center on Deafness Inland Empire team, which comprises people with lived experience as members of and advocates for the D/HH community.

Results: Findings identified a spectrum of needs for mental health apps, including offering American Sign Language and English support, increased education of mental health to reduce stigma around mental health, direct communication with a Deaf worker, and apps that are accessible to a range of community members in terms of culture, resources required, and location.

Conclusions: These findings can inform the development of digital mental health resources and outreach strategies that are appropriate for the D/HH community.

(*JMIR Hum Factors* 2022;9(2):e35641) doi:[10.2196/35641](https://doi.org/10.2196/35641)

KEYWORDS

mental health; deaf and hard of hearing community; mHealth; digital health; needs assessment; deaf; hard of hearing; hearing; focus group; survey; mixed methods; intervention; health app; user needs

Introduction

Accessing mental health services is a challenge in the United States, a challenge that is further magnified for persons who are deaf and hard of hearing (D/HH). D/HH is an umbrella term used to encompass a diverse community. Other terms used by members of the community may include “deaf,” “Deaf,” or “late-deafened.” Following feedback on terminology from our participants, we have chosen to use the term D/HH throughout this paper to refer to this community, and we acknowledge that participants may use different terms to self-identify. Debate exists within the community over Deafness or deafness as disability versus Deafness or deafness as linguistic minority. While this is beyond the scope of this paper, we encourage readers to see Skelton and Valentine for an overview [1]. The D/HH community, which has often been referred to as an “invisible minority” [2], is a community with its own unique culture, traditions, and challenges. Members of the D/HH community may face significant psychosocial challenges and environmental adversity as they navigate an inherently ableist, hearing world [1,3]. Several studies do show that members of the D/HH community experience higher rates of psychological distress [4-6].

When it comes to accessing care, the D/HH community faces significant health care marginalization and health care inequities [3]. D/HH individuals report a lack of availability of mental health services [7], and Critchfield [8] estimates that 80% to 90% of people who are D/HH with severe and persistent mental illness do not receive care. As summarized by Pertz et al [3], this lack of mental health care access is multifaceted and largely stems from systemic barriers facing the community, for example, insurance coverage [9], lack of interpreters for health care visits [10,11], and lack of evidence-based, culturally competent mental health treatment options [12]. Patient outcomes for D/HH persons are better when they receive care from providers who understand Deaf culture, but these are rarely available [13,14]. Pertz et al [3] found that Deaf signers at an integrated medical and behavioral health program with a telemental health (TMH) intervention reported significantly lower depression and anxiety scores from baseline and high satisfaction ratings due to accessible communication and optional ongoing care through a TMH platform. Negative experiences and challenges communicating with ineffective providers can impact treatment engagement and adherence [15], creating general distrust, reluctance, or resistance to the mental health care system [16].

Technology provides opportunities to overcome some of the barriers to accessing mental health care traditionally facing the D/HH community. Many people in the D/HH community report using technology in other aspects of their lives. Examples include text-to-speech apps or smartphone features, such as Ava or Siri; videoconferencing, which is commonly referred to as videophoning in the D/HH community; sound enhancement apps, such as Sound Amplifier; and a variety of visual alert assistive technologies [17]. A national survey by

Maiorana-Basas and Pagliaro [18] suggests that technologies such as texting, emailing, and instant messaging are used at similar rates across the population, regardless of hearing status. The rapid development of technology has led to a proliferation of digital resources designed to support and help people manage their health, and TMH services are effective treatment modalities among the general population (see Langarizadeh et al [19] for a review). TMH services are especially suited in the treatment of D/HH persons because the D/HH community may already have a level of familiarity with visually oriented technologies and assistive technologies, which may help facilitate treatment delivery [20]. TMH may also help facilitate service delivery to D/HH individuals who may otherwise not have a local, culturally competent mental health provider from whom to seek treatment [16]. Furthermore, smartphone access in the United States is increasing, although it is not ubiquitous, and several socioeconomic factors influence access, particularly considering that the average price of a smartphone is now over US \$500 [21]. Those with technology access and digital literacy skills are likely to be younger, highly educated, and possess adequate financial resources. Although technology is often posited as the “great equalizer,” it can also serve to further widen the gaps between privileged and underprivileged groups, who differ in their access to, knowledge of, and ability to make full use of the medium [22].

Mobile health (mHealth) apps for the D/HH community exist, though these largely serve as assistive technologies that aim to augment people’s ability to navigate and communicate in public and with family and connect with other members of the community. In Romero et al’s [23] review of existing mHealth apps for the D/HH community, only two apps from an initial search list of 217 apps were related to mental health. They note that the relatively low yield and high turnover of available apps necessitates more development of apps for the D/HH population. There are no studies, of which we are aware, that have explored mental health apps specifically. Indeed, in our own searches of available resources to inform the development of this study’s methodology, we did not identify any apps to support the mental health of the D/HH community.

In general, while previous studies have identified several challenges among the D/HH community to access mental health resources, it is less understood if and how mental health apps may overcome these challenges. The aims of this study were to explore the mental health needs of the D/HH community and explore how digital resources such as apps may be able to support these needs. To address these aims, we conducted a focus group with 10 community members to get an in-depth understanding of their experiences and perspectives.

Methods

Overview

A community-based participatory approach was used throughout our study to engage community members in multiple stages of

the study. The effort ensured that the data collection content and processes were appropriate, the study design was suitable, and the voices of community members were accurately represented in reporting our findings.

Participants

A total of 10 people participated in one focus group, and 9 of these participants also completed a follow-up survey. Participants were members of the Center on Deafness Inland Empire (CODIE) team and based in Riverside County, California. The CODIE team comprises people with lived experience as members of the D/HH community. CODIE works to advocate for the community by empowering individuals with information, offering training and opportunities, and working to resolve challenges in areas such as communication barriers, peer counseling, independent living skills, community education, and outreach. Participants were invited by email to participate in the focus group by a lead advocate on the CODIE team.

Demographic information was collected using a web-based English-written survey distributed after the focus group. All participants reported comfort with written English, and the survey was developed in partnership with the CODIE team and Riverside County. One participant did not complete the survey, so demographic details describe 9 participants. Given the small sample size, we report the general characteristics of the sample. Participants ranged in age from 30 to 60 years (mean 44.1, SD 11.3). Participants reported their gender as female and identified as White, Black or African American, Asian, American Indian or Alaska Native, or Mexican, or they identified with more than one race. Out of 9 participants, 8 (89%) most often used American Sign Language (ASL) at home, and 7 (78%) participants reported their preferred communication method as ASL.

Measures

For the focus group, the research team developed a focus group facilitator guide with discussion topics and sample questions. Topics and questions were developed and refined based on research partners' interests and their past learnings working with the D/HH community. The research partners consisted of staff from Riverside County Behavioral Health, peer specialists, and the lead advocate on the CODIE team. The research team met with evaluation staff from Riverside County Behavioral Health prior to the focus group to review the questions included in the guide, obtain input on the topics covered, and ensure language used was appropriate and understandable. The lead advocate also provided best practices for facilitating focus group discussions with the D/HH community. First, it was important for the facilitator to have a clear video picture in a well-lit room, tie long hair back, and minimize distractions such as moving objects in the background, so that participants could focus on body language, facial expressions, and lip movements. Second, it was advised for the facilitator to look directly into the camera and speak slowly and clearly to allow for lip reading and interpretation. Third, the facilitator should pause after asking a question to allow for interpretation and look at the interpreter to ensure that interpretation had occurred. Lastly, interpreters should introduce themselves at the beginning of the focus group and provide guidance for participants to pin them on their screen.

While these practices were given specifically for a virtual focus group, many of them are applicable for in-person focus groups too, such as speaking slowly and creating pauses for interpretation (see Balch and Mertens [24] for further lessons learned from D/HH participants on conducting focus groups).

The main focus of this study was on understanding the mental health needs of the D/HH community and how mental health apps may support the D/HH community's mental health needs. Therefore, topics included perspectives on both mental health in general and mental health technologies specifically. Topics covered in the focus group guide included the following:

- Perspectives on mental health within the D/HH community
- Mental health needs and services available for the D/HH community
- Use of and attitudes toward apps and technologies for mental health within the D/HH community
- Challenges and facilitators to using mental health apps and technologies by the D/HH community.

The follow-up written survey was sent 5 days after the focus group. The survey asked additional questions around digital mental health and was intended to supplement findings from the focus group as well as allow participants to express thoughts outside of the focus group setting. The survey questions were developed before the focus group but were refined based on information obtained in the focus group. For example, one survey question asked what aspects of mental health apps were important to participants; the answer options of this question were updated to include certain aspects mentioned during the focus group. The four topics covered in the survey are discussed next.

The first topic was "barriers to mental health resources." Participants were asked to report all barriers, if any, they faced to accessing mental health-related resources. They were instructed to "select all that apply" from a list of options, type free text, or both. The list of barrier options was adapted from the Healthy Minds Study, an annual web-based survey assessing mental health and service use among college students [25].

The second topic was "important aspects about mental health apps." Participants were asked to rate the extent to which different aspects of mental health apps were important to them (eg, "The app is free"). They were asked to rate items on a scale from "not at all important" (1) to "extremely important" (5).

The third topic was "mental health app use." A single question was used to identify whether participants had used mental health apps. In the survey, a mental health app was defined as "an application on your mobile phone or tablet device that helps you manage your mental, emotional, or psychological health or get access to resources to support your mental, emotional, or psychological health." Participants could select whether they had used apps in the past, were currently using apps, had never used apps but would be interested, or had never used apps and were not interested. Participants were also asked to rate three statements related to whether they had the resources required to use mental health apps (eg, "I have the resources necessary to use mental health apps"). The scale ranged from "strongly disagree" (1) to "strongly agree" (5). The items were based on

the facilitating conditions subscale of the unified theory of acceptance and use of technology questionnaire [26]; they were adapted to refer to mental health apps specifically.

The fourth topic was “current and desired resources to support mental health.” Participants were asked to select what resources they currently used and what strategies they wished to use to support their mental health, if any (eg, “informal support, such as talking with or spending time with family or friends”). They were instructed to “select all that apply” from a list of options, type free text, or both.

The complete survey instrument is included in [Multimedia Appendix 1](#).

Procedure

The focus group took place on September 11, 2020, and survey data collection took place between September 16 and 28, 2020. The focus group was held online via Zoom (Zoom Video Communications, Inc), facilitated by hearing research staff, and supported by two interpreters to translate spoken English into ASL and vice versa. Each focus group question was also shown in written English in the chat window of the Zoom session, and participants were able to provide written responses in the chat window. The focus group discussion was audio recorded. The audio recording captured both the ASL translated to spoken English and the chat messages, which were read aloud. The duration of the focus group was 2 hours. The survey was distributed via Qualtrics and took approximately 20 minutes to complete. Participants received a US \$30 gift card for their participation in the focus group and a US \$10 gift card for completing the survey.

Ethical Considerations

The study was approved by the University of California, Irvine, Institutional Review Board (IRB; review number No. 20195406). Prior to the focus group, participants were emailed a study information sheet that was reviewed and approved by the CODIE lead advocate and the IRB. The sheet was then reviewed in the focus group session, with an opportunity for participants to ask questions. Participants were asked for their permission to audio record the conversation at the start of the focus group.

Analysis

The audio recording of the focus group was transcribed. The analytical framework used to analyze the transcript was the six-phased approach of thematic analysis as described in Braun and Clarke [27], which involves the following: (1) familiarizing yourself with the data, (2) generating initial codes, (3) searching for themes, (4) reviewing themes, (5) defining and naming themes, and (6) writing up the analysis. We adopted an interpretivist epistemological position and used an inductive analysis approach: there was no pre-existing coding scheme, and codes were created based on what emerged from the data. The qualitative analysis software Atlas.ti (version 22.0.1; Scientific Software Development GmbH) was used to code the transcript. The initial coding (phases 1-3) was done by one of the non-D/HH researchers, who is a trained PhD researcher with expertise in user experience and thematic analysis. For phase

4, a preliminary summary of findings was shared with study participants and other members of the D/HH community and research team who attended the focus group in order to check that this was what was said, to corroborate, to correct or extend interpretations of findings, and to further refine themes. These findings were discussed over email and during a video call meeting, where an interpreter was present to support the discussion. For phase 5, themes were defined and named by the non-D/HH research team. For phase 6, a draft of the write-up was shared with members of the D/HH community and research team who attended the focus group to provide feedback, craft the language, and add details.

We analyzed the survey data using descriptive statistics in the form of the number of people who selected certain answers. The statistical software SAS (version 9.4; SAS Institute Inc) [28] was used for analysis of the survey data. The main purpose of the survey was to supplement findings from the focus group and describe the study sample (eg, demographic information, the number of participants who had used mental health apps before, and the number of participants who wanted to use certain mental health app features that were discussed during the focus group).

Results

The following section presents an overview of study results. Unless otherwise specified, results are based on the focus group. Illustrative quotes are provided for each theme.

Current and Desired Strategies to Support Mental Health

Participants had not used any mental health apps before. A total of 6 out of 9 (67%) participants indicated on the survey that they were interested in using one, and 3 (33%) participants indicated they were not interested. Though mental health apps were not commonly used, participants shared that they used other online resources to support mental health, such as spiritual classes, meditation, and ASL yoga:

I use deaf spirituality. That kind of covers a lot of different things. Healing, holistic healing. There's meditation. I use a website as well as a good resource, oh, to find practitioners who use sign language for all of those types of things.

Other apps that were reported to be used were communication apps to connect with others, such as WhatsApp, Zoom, and Skype.

Participants were asked on the survey what their current and desired strategies were to support their mental health. The most common strategies currently used involved informal support connecting with friends or family (n=6, 67%), peer support (n=6, 67%), and use of social media (n=5, 56%). The most common desired resource was professional mental health services (n=6, 67%), followed by activities like writing, painting, and playing or making music (n=5, 56%); online forums or communities (n=4, 44%); and websites (n=4, 44%). A total of 3 (33%) participants reported wanting to use online chat and peer support, as well as exercise programs or activities to manage their mental health.

Participants were also asked on the survey about what the broad D/HH community would like to be able to do with mental health resources. They felt the D/HH community would be most interested in resources that allow them to talk with other people to give and get support (n=6, 67%) as well as those that allow them to express themselves or have an outlet through art, photos, or writing (n=6, 67%). A total of 5 (56%) participants also reported several other possible interests, such as identifying or recognizing symptoms, working through negative emotions and thoughts, connecting with a professional, and getting information about how to cope with stress, grief and loss, trauma, and relationship issues.

Challenges to Using Mental Health Apps

Support for a Spectrum of Language and Linguistic Needs Within the Community

The most reported barrier to accessing mental health services on the survey and during the focus group was the difficulty in finding mental health care providers that knew ASL. Similarly, the main barrier to using online mental health tools specifically, as reported on the survey, was difficulty finding a tool that supported ASL.

Beyond a lack of ASL support, participants reported issues to accessing mental health services with respect to communication, access, and feeling welcome. Participants shared that there are a range of language and linguistic needs within the community, with some people feeling more comfortable with English, whereas others are more comfortable with ASL. Furthermore, there are different literacy levels within the community in terms of understanding English. Participants recommended providing different options to present content through a digital mental health intervention, such as text, videos, and icons, and providing ASL video where possible. One participant noted the following:

Talking about English and ASL, there's neither one that is better than the other. It's a matter of what the person feels most comfortable with...You also don't want it to be just ASL only, it might force somebody out of their comfort zone. So we need to consider that spectrum of language and linguistic needs and comfort levels, which is really wide.

In the context of differing linguistic needs, participants also discussed knowledge gaps in relation to mental health concerns. The D/HH community misses incidental learning opportunities around mental health, which happen when people gain knowledge from informal interactions and overhearing conversations that can be related to societal changes in attitude toward mental health. These learning opportunities typically rely on spoken language.

Lack of Accessible Services

Even if an interpreter could be provided to aid communication with a mental health provider, participants reported that many community members would not feel comfortable with having an interpreter present and would feel safer if they could speak directly to a mental health provider with the same language. One participant explained the following:

Sometimes interpreters will use different word choices and it's not what I mean. Or you know, confidentiality, because people may want to keep that privacy.

One participant mentioned that to build a connection with a health provider, it helps to talk to someone who looks and signs like them. An additional barrier to accessing mental health services was that providers were not sensitive enough to cultural differences. For example, a participant explained that hearing providers do not have a “deaf heart” and the sensitivity or the same experiences as them.

Participants said there was a lack of Deaf workers in the mental health profession and that it was challenging to find mental health services for the D/HH community. Participants expressed concerns that there was a lack of accessible resources overall and for specific services, such as marriage counseling, anger management, substance abuse treatment, and support for domestic violence. For services that were available, participants said that community members were sometimes limited in terms of their insurance and what services they could access. For example, one participant commented that services may only be available out of their state and, thus, not covered by insurance, or that the only ASL services available are very basic.

Stigma Around Mental Health

There was a consensus during the focus group that stigma around mental health was a considerable challenge in the community, and that community members did not want other people to know they were accessing mental health services. Participants had concerns over the use of the term “mental health” and said that positive and uplifting terms centered around spirituality and healing would be more appropriate, stating that these would resonate more with the community and signal that a positive experience is forthcoming. For example, one participant suggested the name “Healing Hands” for a mental health app.

Participants expressed concerns that for many community members, miscommunication has had negative repercussions in the past, which can increase stigma, and there are fears of experiencing negative side effects of getting treatment. For example, members with children may have fears that if they talk about their mental health challenges, their family, such as their children, may be taken away or there may be financial consequences. One participant stated that mental health can be perceived as “just another thing wrong with my head.”

Participants noted that community members may have privacy concerns around the use of mental health services, such as concerns around whether their information was going to be safe. On the survey, 3 (33%) participants indicated that they had privacy concerns on their personal information being visible by using mental health apps.

Participants expressed a need for increased education and awareness around mental health, and to promote a message that mental health services are helpful in a good way, that it was okay to seek help, and that mental health is for everyone:

Nowadays what I'm seeing is Deaf and Hard of Hearing people putting vlogs emphasizing it's okay

to feel whatever you're feeling, and it's okay to look for help. And I think that's key, if developing an app...to emphasize that. That's what most of society is doing at this moment.

Facilitators to Engagement With Mental Health Apps

Overview

Participants gave several recommendations on marketing a mental health app to the D/HH community. Examples included using posters and signs, scrolling and video advertisements for the app at medical offices and social service offices, contacting nonprofit organizations that service the D/HH community, and word of mouth. It was important that the marketing materials supported a feeling of being welcome, for example, through visual advertisements that showed the step-by-step process of using the app. Participants preferred the app to be advertised with ASL people signing, using more visuals than words. Participants pointed out that instead of an “interpreter” sign, a better sign would be the two-handed sign for “peer,” “advocate,” or “support,” ideally with hands of different colors and genders. They also approved of the “same same” sign.

With respect to speaking with a health provider, participants expressed a need to be able to choose a specific person with whom they felt comfortable talking. Some people may have experienced trauma with past providers and wanted to talk to someone who would be a good fit for them regarding language and other characteristics, like gender. It was important to have diversity within the community and presented on an app, in order to make the app accessible to everyone.

Lastly, if a mental health app were to be developed that was inclusive of the D/HH community, participants expressed a preference for an app that would be useful to everyone, not just members of the D/HH community. Clicking on an app that would be specifically labeled for the D/HH community could give a feeling of being singled out. As one participant explained:

We want to try and keep that general to have access to things instead of feeling like, ‘oh okay, I have to click on this because it says deaf,’ that singles me out.

Immediate and Continuous Access to Resources

Participants placed importance on the fact that a mental health app should be accessible to a range of people in terms of language, culture, resources required to use the app, and location. They explained that community members may have limited data or memory on their phone, no access to high-speed internet, or no access to a computer. A mobile app was the preferred platform to enable people to access resources on the go. On the survey, 5 (56%) participants indicated that they had concerns about their mobile data plan when using their mobile device, 4 (44%) participants did not have the necessary resources to use mental health apps, and 3 (33%) participants were concerned about having enough space to download apps on their smartphone.

Participants also expressed value for an app to provide immediate access to resources and services. At the time of the focus group, people had to go through a long intake process before they could connect to mental health services:

Having that immediate assistance, to somebody live or whatever it is, right there is important rather than having to go through all of these different things...you have to go through all of that demographic information and you have to basically tell your life story before you can get to somebody.

Given the range of literacy levels within the D/HH community, participants worried that people may not understand all intake questions, which can slow down the process further and reduce interest in engagement. A participant mentioned that having the intake available in both English and ASL would likely make the intake process go smoother.

Participants also said that it would be ideal to have unlimited access to resources, as opposed to there being a limit to the number of times they could access them. Especially during the global COVID-19 pandemic, services were sometimes used by people just to connect and talk to someone, and some consumers accessed mental health services multiple times a day.

Though participants were part of a local community-based agency, they named several benefits of making an app globally accessible to anyone, rather than tying it to a specific location. First, people may be located elsewhere but prefer coming to a specific organization, such as CODIE, for services. Second, there were benefits to working together with other organizations. If someone is in need outside of standard business hours, there may not be anyone near them to help, but there may be someone awake in another time zone and location who can provide support. Third, these collaborations could facilitate linking to resources from other organizations, in order to spread awareness and help the broader D/HH community. Participants suggested making apps available to people from Gallaudet University, the only university in the world where students learn in ASL and English.

Important Features for Mental Health Apps

Participants were asked on the survey what the most important aspects were that mental health apps should offer. The most important aspects were suicide prevention support, emergency support, peer support and chat, and telehealth (ie, referring to a direct connection to clinical mental health services within the app). A chatbot was rated as the least important.

During the focus group, participants reported that it was important for people to interact with a human and not an avatar. An avatar is a computerized figure, such as an icon, that can represent a simplified figure of a person. The term avatar can refer to different things and can also be used for digitally created characters that turn speech into sign language. Participants' reservations about an avatar were that it lacked facial expressions and body language, which are important to have in addition to signs in communication. One participant explained as follows:

We want an actual person, not an avatar because an avatar lacks that body language, those expressions that, a lot of people that maybe are at the gestural language level would understand the body language more than the actual signs.

Furthermore, they reported that it was valuable to build a real relationship with a human being. It was not discussed during the focus group whether participants did not want avatars at all or whether avatars could be used to supplement interaction in an app, if it also allowed connection with a live person.

Participants mentioned past use of Deaf clubs, which are places where Deaf people can meet face-to-face and socialize [29]. Participants shared that they used Deaf clubs for socializing, fun, and games, and expressed the desire to see these Deaf clubs being used for mental health support. It was discussed whether an app could have something similar to a Deaf club.

Inclusion of Diverse Community Members in Technology Development

To reduce stigma around seeking help for mental health, participants suggested having members of the community contribute to an app, for example, through blogs or short videos to share their experiences and knowledge. Participants liked the aspect of inclusivity where visitors can be “part of the news”:

Like a blog that people could add—share their experiences and their knowledge and their education...They could be a part of the news basically. Instead of watching the news with captions, they could watch this person signing that.

In addition to providing content, participants recommended having community members involved in the design process to give feedback on features and on what can be improved. Some participants reported that it would also be valuable to have community members provide guidance on how to access and use the app, for example, through instruction videos and visual posters with step-by-step instructions.

Discussion

Principal Findings

The aim of this paper was to understand the mental health needs of the D/HH community and how mental health apps may be able to support these needs. In line with previous work [7,9], participants indicated that community members had limited access to mental health resources. Digital solutions, such as mental health apps, may increase access to resources, but our study highlighted that it is important to take certain factors into consideration to facilitate engagement with such apps.

Some of the themes we found are common barriers among other mental health-seeking populations as well, such as stigma [30]. This barrier may be exacerbated in the D/HH community due to missed incidental learning opportunities about mental health. Furthermore, similar to previous work with hearing populations [31], participants valued immediate access to resources. Participants rated access to suicide prevention support and peer chat as one of the most important features to include in a mental health app, which resonates with work with other communities: a recent study with essential workers found that one of the most desired features for a mental health technology was the ability to chat with a mental health professional or peer, and a link to mental health resources and crisis support [31].

Factors that may be more unique to the D/HH community are the need for both ASL and English support, and the finding that participants wanted a general app that is inclusive of the D/HH community, rather than an app exclusively made for them. For example, participants emphasized the importance of including members of the D/HH community on the app, but to market an app as usable for everyone to avoid singling out the D/HH community. This finding further supports the need for customization and personalization of mental health apps [32,33] and the importance of inclusive design and designing for a wider population [34]. The ability to customize an app to a user's personal needs can facilitate feelings of perceived fit to a user's culture and values [35], without singling out a particular community.

While further follow-up studies are recommended to corroborate themes with a larger population, initial takeaways can be extracted from our findings to inform creation and development of digital mental health resources, such as apps. Below we outline several learnings that may be important to consider when developing digital mental health resources for the D/HH population.

Support for a Spectrum of Language and Linguistic Needs Within the Community

Similar to previous work [10,36], the greatest barrier to accessing mental health services identified by participants pertained to communication issues. Participants reported a lack of Deaf workers and mental health care providers that knew ASL. It is, thus, important to support a spectrum of linguistic needs within the community.

Participants in our focus group primarily highlighted the limitations of English-language mental health services, rather than positive experiences. A previous focus group with D/HH community members indicated that there may be social pressure during ASL focus groups that limits participants from sharing any positive experiences with English health care communication [37]. While our study participants acknowledged that some community members may prefer English, the main issue was that there needs to be support for a *range* of linguistic needs, rather than English or ASL support alone. Participants shared that there are various language and linguistic needs within the community, with some people feeling more comfortable with English and ASL, and there are different literacy levels in terms of understanding English. Our survey further found that a lack of Deaf workers in the mental health profession was not the only barrier to accessing resources, but it was the most common.

While previous studies found difficulties in accessing mental health resources due to a lack of interpreters [37,38], participants in our study indicated that the availability of an interpreter alone may not be sufficient. Even with an interpreter present, participants explained that community members may not feel comfortable talking via an interpreter and prefer to communicate with a Deaf worker directly. This sentiment is consistent with findings from Steinberg et al [39] who found a preference for health care practitioners who are fluent in ASL and support from other Deaf individuals.

To support a range of linguistic needs, participants recommended providing different options for presenting content through an app, such as text, videos, and icons, and providing ASL video where possible. Participants also provided marketing suggestions to support a feeling of being welcome, for example, through visual advertisements that show the step-by-step process of using the app.

Stigma and Appropriate Use of Terminology

Similar to prior studies with D/HH community members [39], study participants expressed that there was stigma within the community around mental health issues and seeking help for these issues. Though stigma can be a common barrier among the population in general [30], it may especially be an issue for the D/HH population, as they are not exposed to mental health issues and information in the same way as the general public [38]. Even though study participants expressed a need for more Deaf workers in the mental health profession, previous work found that because of this lack of exposure, Deaf workers may be less knowledgeable about mental health issues [38].

Prior work on mental health apps has suggested that delivering support through technology can overcome stigma barriers, as people do not need to know one is seeking help [40]. However, participants in our study still had concerns about their information not being private through an app. To help mitigate privacy concerns, it is important to be transparent on how app data are collected and stored and how they will be used. Furthermore, participants recommended that instead of using the term mental health, positive and uplifting terms around healing are preferred in order to facilitate adoption of a mental health app.

Education Around Mental Health

Communication issues can complicate accurate reporting of mental health prevalence in the D/HH community [41,42], and our findings further showed that D/HH members can often miss out on informal conversations and may not be as knowledgeable about mental health as the hearing population. This finding highlights that increased education around mental health may be especially important for this community. Participants expressed a need for increased education and awareness around mental health, for example, through short videos and by having members of the community share their experiences. Participants stated that there was a need for people to understand that mental health services can be helpful, and that strong mental health is a goal for everyone.

Include Members of the D/HH Community and Market for Broader Community

Ideally, participants preferred to have direct communication with a Deaf worker that had the sensitivity and experience to communicate with members of the community. Participants

also recommended involving community members in providing content and sharing feedback about improving app features. It was important to have an app that is inclusive of, but not exclusively for, the D/HH community. Participants preferred an app that would be useful for anyone and that would not just be focused on their community, which may exacerbate feelings of being singled out.

Limitations and Future Work

The study has several limitations. First, care should be taken to generalize its findings to the broader community. An advantage of a focus group setting is that it has been shown to be a suitable methodology for Deaf culture to gather and share information in a safe setting [37,38], but sample size is limited. In addition, participants in the focus group were engaged with an advocacy group and involved in the community, so their experiences may be different than those of the general community. Participants had experienced hearing loss since birth or early in life, and their experiences may differ from those who experience hearing loss because of old age or those who experience hearing loss later in life. Our study offers insights into how mental health needs of the D/HH community may be supported through digital therapeutics, which would be worthwhile to explore further in a larger-scale study. Second, an English written survey was used to collect participants' demographic information. Though all participants were able to read and write in English, the majority of participants' preferred language was ASL, and one participant did not complete the survey. Our study results have since been used to inform a collaborative effort to create an ASL survey for broader needs assessment with the D/HH community. Third, results were collected during the COVID-19 pandemic, which may have increased mental health concerns and interest in mental health resources. Lastly, we used the English translation of the focus group discussion for data analysis. There may be limitations in using an English translation, as information may be filtered and expressions can differ from ASL. For example, personal pronouns in ASL are not gender specific [37]. To ensure that our analysis and findings accurately represented participants' views, we refined our findings through member checking with the focus group participants who are members of the D/HH community.

Conclusions

This study looked at the mental health needs of the D/HH community and how mental health apps may be able to support these needs. There was a need for more Deaf workers and ASL support to support a spectrum of linguistic needs; a need for increased education to reduce stigma around mental health; a need for an app that is accessible to a range of people in terms of culture, resources required, and location; and a need for immediate and unlimited access to resources. These findings are important to consider for the development and dissemination of mental health apps to meet the needs of the D/HH community.

Acknowledgments

We deeply appreciate the members of CODIE who shared their experiences and feedback with us and were patient as we learned more about the D/HH community. Without their expertise, these insights would not be possible. We would also like to thank the

interpreters as well as Suzanna Juarez-Williamson, Riverside County, the Peer Support Team of the Help@Hand Project, and the Help@Hand Evaluation Team.

This work was funded by Help@Hand (contract No. 417-ITSUCI-2019), a project overseen by the California Mental Health Service Authority (CalMHSA). CalMHSA reviewed the manuscript for confidentiality. The information or content and conclusions presented here are those of the authors and should not be construed as official position or policy of, nor should any endorsements be inferred by, the participating Help@Hand counties or CalMHSA.

Conflicts of Interest

SMS has received consulting payments from Otsuka Pharmaceuticals and Trusst (K Health) and is a member of the Headspace Scientific Advisory Board, for which he receives compensation.

Multimedia Appendix 1

Survey items.

[[DOCX File, 41 KB - humanfactors_v9i2e35641_app1.docx](#)]

References

1. Skelton T, Valentine G. 'It feels like being Deaf is normal': An exploration into the complexities of defining D/deafness and young D/deaf people's identities. *Can Geogr* 2003 Dec 08;47(4):451-466. [doi: [10.1111/j.0008-3658.2003.00035.x](#)]
2. Hunt G. Trapped in the shadows: The working lives of invisible minorities. In: Pupo N, Glenday D, Duffy A, editors. *The Shifting Landscape of Work*. Toronto, ON: Nelson Education Ltd; 2010:101-118.
3. Pertz L, Plegue M, Diehl K, Zazove P, McKee M. Addressing mental health needs for deaf patients through an integrated health care model. *J Deaf Stud Deaf Educ Internet* 2018 Jul 01;23(3):240-248. [doi: [10.1093/deafed/eny002](#)] [Medline: [29562357](#)]
4. Bridgman G, MacPherson B, Rako M, Campbell J, Manning V, Norman-Kelly T. A national epidemiological survey of mental illness in the New Zealand Deaf community. In: *Proceedings of the 5th European and 2nd World Conference on Mental Health and Deafness*. 2000 Presented at: The 5th European and 2nd World Conference on Mental Health and Deafness; October 18-21, 2000; Copenhagen, Denmark p. 216-234.
5. Fellinger J, Holzinger D, Dobner U, Gerich J, Lehner R, Lenz G, et al. Mental distress and quality of life in a deaf population. *Soc Psychiatry Psychiatr Epidemiol* 2005 Sep;40(9):737-742. [doi: [10.1007/s00127-005-0936-8](#)] [Medline: [16143834](#)]
6. de Graaf R, Bijl RV. Determinants of mental distress in adults with a severe auditory impairment: Differences between prelingual and postlingual deafness. *Psychosom Med* 2002;64(1):61-70. [doi: [10.1097/00006842-200201000-00009](#)] [Medline: [11818587](#)]
7. Feldman DM, Gum A. Multigenerational perceptions of mental health services of deaf adults in Florida. *American Annals of the Deaf* 2007;152(4):391-397. [doi: [10.1353/aad.2008.0001](#)] [Medline: [18257508](#)]
8. Critchfield AB. Meeting the Mental Health Needs of Persons Who Are Deaf. Alexandria, VA: National Technical Assistance Center for State Mental Health Planning (NTAC); 2002 May. URL: https://mh.alabama.gov/wp-content/uploads/2019/01/ODS_MeetingTheMentalHealthNeedsOfPersonsWhoAreDeaf.pdf [accessed 2021-12-11]
9. McKee MM, McKee K, Winters P, Sutter E, Pearson T. Higher educational attainment but not higher income is protective for cardiovascular risk in Deaf American Sign Language (ASL) users. *Disabil Health J* 2014 Jan;7(1):49-55 [FREE Full text] [doi: [10.1016/j.dhjo.2013.07.001](#)] [Medline: [24411507](#)]
10. Alexander A, Ladd P, Powell S. Deafness might damage your health. *Lancet* 2012 Mar 17;379(9820):979-981. [doi: [10.1016/S0140-6736\(11\)61670-X](#)] [Medline: [22423872](#)]
11. Mitchell J, Howell C, Turnbull D, Murphy M. Computer-assisted group therapy for the treatment of depression and anxiety in general practice. *Prim Care Ment Health* 2005;3(1):27-39.
12. Kushalnagar P, Holcomb J, Sadler G. Genetic testing and eHealth usage among Deaf women. *J Genet Couns* 2019 Oct;28(5):933-939 [FREE Full text] [doi: [10.1002/jgc4.1134](#)] [Medline: [31180177](#)]
13. Pollard RQ, Betts WR, Carroll JK, Waxmonsky JA, Barnett S, deGruy FV, et al. Integrating primary care and behavioral health with four special populations: Children with special needs, people with serious mental illness, refugees, and deaf people. *Am Psychol* 2014;69(4):377-387. [doi: [10.1037/a0036220](#)] [Medline: [24820687](#)]
14. Landsberger S, Sajid A, Schmelkin L, Diaz D, Weiler C. Assessment and treatment of deaf adults with psychiatric disorders: A review of the literature for practitioners. *J Psychiatr Pract* 2013 Mar;19(2):87-97. [doi: [10.1097/01.pra.0000428555.48588.f9](#)] [Medline: [23507810](#)]
15. Sudore RL, Landefeld CS, Pérez-Stable EJ, Bibbins-Domingo K, Williams BA, Schillinger D. Unraveling the relationship between literacy, language proficiency, and patient-physician communication. *Patient Educ Couns* 2009 Jun;75(3):398-402 [FREE Full text] [doi: [10.1016/j.pec.2009.02.019](#)] [Medline: [19442478](#)]
16. Crowe TV. Is telemental health services a viable alternative to traditional psychotherapy for deaf individuals? *Community Ment Health J* 2017 Feb;53(2):154-162. [doi: [10.1007/s10597-016-0025-3](#)] [Medline: [27260308](#)]

17. Wilson JAB, Schild S. Provision of mental health care services to deaf individuals using telehealth. *Prof Psychol Res Pr* 2014;45(5):324-331. [doi: [10.1037/a0036811](https://doi.org/10.1037/a0036811)]
18. Maiorana-Basas M, Pagliaro CM. Technology use among adults who are deaf and hard of hearing: A national survey. *J Deaf Stud Deaf Educ* 2014 Jul;19(3):400-410. [doi: [10.1093/deafed/enu005](https://doi.org/10.1093/deafed/enu005)] [Medline: [24661981](https://pubmed.ncbi.nlm.nih.gov/24661981/)]
19. Langarizadeh M, Tabatabaei M, Tavakol K, Naghipour M, Rostami A, Moghbeli F. Telemental health care, an effective alternative to conventional mental care: A systematic review. *Acta Inform Med* 2017 Dec;25(4):240-246 [FREE Full text] [doi: [10.5455/aim.2017.25.240-246](https://doi.org/10.5455/aim.2017.25.240-246)] [Medline: [29284913](https://pubmed.ncbi.nlm.nih.gov/29284913/)]
20. Behl DD, Kahn G. Provider perspectives on telepractice for serving families of children who are deaf or hard of hearing. *Int J Telerehab* 2015 Jul 29;3-12. [doi: [10.5195/jjt.2015.6170](https://doi.org/10.5195/jjt.2015.6170)]
21. O'Dea S. Smartphone (consumer/enterprise) average price in the United States 2014-2025. Statista. 2022. URL: <https://www.statista.com/statistics/619830/smartphone-average-price-in-the-us/> [accessed 2022-04-06]
22. Bitman N, John N. Deaf and hard of hearing smartphone users: Intersectionality and the penetration of ableist communication norms. *J Comput Mediated Commun* 2019;24(2):56-72. [doi: [10.1093/jcmc/zmy024](https://doi.org/10.1093/jcmc/zmy024)]
23. Romero RL, Kates F, Hart M, Ojeda A, Meirum I, Hardy S. Quality of deaf and hard-of-hearing mobile apps: Evaluation using the Mobile App Rating Scale (MARS) with additional criteria from a content expert. *JMIR Mhealth Uhealth* 2019 Oct 30;7(10):e14198 [FREE Full text] [doi: [10.2196/14198](https://doi.org/10.2196/14198)] [Medline: [31670695](https://pubmed.ncbi.nlm.nih.gov/31670695/)]
24. Balch G, Mertens D. Focus group design and group dynamics: Lessons from deaf and hard of hearing participants. *Am J Eval* 1999;20(2):265-277. [doi: [10.1016/s1098-2140\(99\)00019-3](https://doi.org/10.1016/s1098-2140(99)00019-3)]
25. Lipson SK, Kern A, Eisenberg D, Breland-Noble AM. Mental health disparities among college students of color. *J Adolesc Health* 2018 Sep;63(3):348-356. [doi: [10.1016/j.jadohealth.2018.04.014](https://doi.org/10.1016/j.jadohealth.2018.04.014)] [Medline: [30237000](https://pubmed.ncbi.nlm.nih.gov/30237000/)]
26. Venkatesh V, Thong JYL, Xu X. Consumer acceptance and use of information technology: Extending the unified theory of acceptance and use of technology. *MIS Q* 2012;36(1):157-178. [doi: [10.2307/41410412](https://doi.org/10.2307/41410412)]
27. Braun A, Clarke V. Using thematic analysis in psychology. *Qual Res Psychol* 2006 Aug;3(2):77-101. [doi: [10.1002/0471142905.hg1201s50](https://doi.org/10.1002/0471142905.hg1201s50)] [Medline: [18428388](https://pubmed.ncbi.nlm.nih.gov/18428388/)]
28. SAS® 9.4 — Today and Tomorrow. SAS Support. Cary, NC: SAS Institute Inc; 2013. URL: <https://support.sas.com/software/94/> [accessed 2022-04-05]
29. Padden C. The decline of Deaf clubs in the US: A treatise on the problem of place. In: Bauman D, editor. *Sightings: Explorations in Deaf Studies*. Minneapolis, MN: University of Minnesota Press; 2007.
30. Gulliver A, Griffiths KM, Christensen H. Perceived barriers and facilitators to mental health help-seeking in young people: A systematic review. *BMC Psychiatry* 2010 Dec 30;10(1):113 [FREE Full text] [doi: [10.1186/1471-244x-10-113](https://doi.org/10.1186/1471-244x-10-113)]
31. Mata-Greve F, Johnson M, Pullmann M, Friedman E, Griffith Filippo I, Comtois K, et al. Mental health and the perceived usability of digital mental health tools among essential workers and people unemployed due to COVID-19: Cross-sectional survey study. *JMIR Ment Health* 2021 Aug 05;8(8):e28360 [FREE Full text] [doi: [10.2196/28360](https://doi.org/10.2196/28360)] [Medline: [34081592](https://pubmed.ncbi.nlm.nih.gov/34081592/)]
32. Strecher V, McClure J, Alexander G, Chakraborty B, Nair V, Konkel J, et al. The role of engagement in a tailored web-based smoking cessation program: Randomized controlled trial. *J Med Internet Res* 2008 Nov 04;10(5):e36 [FREE Full text] [doi: [10.2196/jmir.1002](https://doi.org/10.2196/jmir.1002)] [Medline: [18984557](https://pubmed.ncbi.nlm.nih.gov/18984557/)]
33. Torous J, Nicholas J, Larsen ME, Firth J, Christensen H. Clinical review of user engagement with mental health smartphone apps: Evidence, theory and improvements. *Evid Based Ment Health* 2018 Aug;21(3):116-119. [doi: [10.1136/eb-2018-102891](https://doi.org/10.1136/eb-2018-102891)] [Medline: [29871870](https://pubmed.ncbi.nlm.nih.gov/29871870/)]
34. Clarkson P, Coleman R, Keates S, Lebbon C, editors. *Inclusive Design: Design for the Whole Population*. London, UK: Springer Science & Business Media; 2003.
35. Borghouts J, Eikley E, Mark G, De Leon C, Schueller SM, Schneider M, et al. Barriers to and facilitators of user engagement with digital mental health interventions: Systematic review. *J Med Internet Res* 2021 Mar 24;23(3):e24387 [FREE Full text] [doi: [10.2196/24387](https://doi.org/10.2196/24387)] [Medline: [33759801](https://pubmed.ncbi.nlm.nih.gov/33759801/)]
36. Steinberg A. Issues in providing mental health services to hearing-impaired persons. *Hosp Community Psychiatry* 1991 Apr;42(4):380-388. [doi: [10.1176/ps.42.4.380](https://doi.org/10.1176/ps.42.4.380)] [Medline: [2050351](https://pubmed.ncbi.nlm.nih.gov/2050351/)]
37. Steinberg AG, Barnett S, Meador HE, Wiggins EA, Zazove P. Health care system accessibility. Experiences and perceptions of deaf people. *J Gen Intern Med* 2006 Mar;21(3):260-266 [FREE Full text] [doi: [10.1111/j.1525-1497.2006.00340.x](https://doi.org/10.1111/j.1525-1497.2006.00340.x)] [Medline: [16499543](https://pubmed.ncbi.nlm.nih.gov/16499543/)]
38. Cabral L, Muhr K, Savageau J. Perspectives of people who are deaf and hard of hearing on mental health, recovery, and peer support. *Community Ment Health J* 2013 Dec;49(6):649-657. [doi: [10.1007/s10597-012-9569-z](https://doi.org/10.1007/s10597-012-9569-z)] [Medline: [23149648](https://pubmed.ncbi.nlm.nih.gov/23149648/)]
39. Steinberg AG, Sullivan VJ, Loew RC. Cultural and linguistic barriers to mental health service access: The deaf consumer's perspective. *Am J Psychiatry* 1998 Jul;155(7):982-984. [doi: [10.1176/ajp.155.7.982](https://doi.org/10.1176/ajp.155.7.982)] [Medline: [9659872](https://pubmed.ncbi.nlm.nih.gov/9659872/)]
40. Kern A, Hong V, Song J, Lipson SK, Eisenberg D. Mental health apps in a college setting: Openness, usage, and attitudes. *Mhealth* 2018;4:20 [FREE Full text] [doi: [10.21037/mhealth.2018.06.01](https://doi.org/10.21037/mhealth.2018.06.01)] [Medline: [30050916](https://pubmed.ncbi.nlm.nih.gov/30050916/)]
41. Diaz DR, Landsberger SA, Povlinski J, Sheward J, Sculley C. Psychiatric disorder prevalence among deaf and hard-of-hearing outpatients. *Compr Psychiatry* 2013 Oct;54(7):991-995. [doi: [10.1016/j.comppsy.2013.04.004](https://doi.org/10.1016/j.comppsy.2013.04.004)] [Medline: [23688857](https://pubmed.ncbi.nlm.nih.gov/23688857/)]
42. Black PA, Glickman NS. Demographics, psychiatric diagnoses, and other characteristics of North American Deaf and hard-of-hearing inpatients. *J Deaf Stud Deaf Educ* 2006;11(3):303-321. [doi: [10.1093/deafed/enj042](https://doi.org/10.1093/deafed/enj042)] [Medline: [16687730](https://pubmed.ncbi.nlm.nih.gov/16687730/)]

Abbreviations

ASL: American Sign Language
CalMHSA: California Mental Health Service Authority
CODIE: Center on Deafness Inland Empire
D/HH: deaf and hard of hearing
IRB: Institutional Review Board
mHealth: mobile health
TMH: telemental health

Edited by A Kushniruk; submitted 11.12.21; peer-reviewed by R Ramadurai, L Welch; comments to author 01.02.22; revised version received 14.02.22; accepted 21.02.22; published 11.04.22.

Please cite as:

Borghouts J, Neary M, Palomares K, De Leon C, Schueller SM, Schneider M, Stadnick N, Mukamel DB, Sorkin DH, Brown D, McCleerey-Hooper S, Moriarty G, Eikey EV

Understanding the Potential of Mental Health Apps to Address Mental Health Needs of the Deaf and Hard of Hearing Community: Mixed Methods Study

JMIR Hum Factors 2022;9(2):e35641

URL: <https://humanfactors.jmir.org/2022/2/e35641>

doi: [10.2196/35641](https://doi.org/10.2196/35641)

PMID: [35404259](https://pubmed.ncbi.nlm.nih.gov/35404259/)

©Judith Borghouts, Martha Neary, Kristina Palomares, Cinthia De Leon, Stephen M Schueller, Margaret Schneider, Nicole Stadnick, Dana B Mukamel, Dara H Sorkin, Dakota Brown, Shannon McCleerey-Hooper, Gloria Moriarty, Elizabeth V Eikey. Originally published in JMIR Human Factors (<https://humanfactors.jmir.org>), 11.04.2022. This is an open-access article distributed under the terms of the Creative Commons Attribution License (<https://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Human Factors, is properly cited. The complete bibliographic information, a link to the original publication on <https://humanfactors.jmir.org>, as well as this copyright and license information must be included.

Original Paper

Process and Information Needs When Searching for and Selecting Apps for Smoking Cessation: Qualitative Study Using Contextual Inquiry

Ylva Hendriks¹, MSc, MPIM; Sebastiaan Peek¹, PhD; Maurits Kaptein², PDEng, PhD; Inge Bongers^{1,3}, PhD

¹Tranzo, Tilburg School of Social and Behavioral Sciences, Tilburg University, Tilburg, Netherlands

²Jheronimus Academy of Data Science, 's-Hertogenbosch, Netherlands

³Research Unit Evidence Based Management of Innovation, Mental Health Care Institute Eindhoven, Eindhoven, Netherlands

Corresponding Author:

Ylva Hendriks, MSc, MPIM

Tranzo, Tilburg School of Social and Behavioral Sciences

Tilburg University

PO Box 90153

Tilburg, 5000 LE

Netherlands

Phone: 31 13 466 4892

Email: ylva.hendriks@tilburguniversity.edu

Abstract

Background: Hundreds of apps are available to support people in their quest to quit smoking. It has been hypothesized that selecting an app from a sizable volume without any aid can be overwhelming and difficult. However, little is known about how people choose apps for smoking cessation and what exactly people want to know about an app before choosing to install it. Understanding the decision-making process may ultimately be helpful in creating tools to help people meaningfully select apps.

Objective: The aim of this study is to obtain insights into the process of searching and selecting mobile apps for smoking cessation and map the range of actions and the accompanying reasons during the search, focusing on the information needs and experiences of those who aim to find an app.

Methods: Contextual inquiries were conducted with 10 Dutch adults wanting to quit smoking by using an app. During the inquiries, we observed people as they chose an app. In addition, 2 weeks later, there was a short semistructured follow-up interview over the phone. Through convenience and purposive sampling, we included participants differing in gender, age, and educational level. We used thematic analysis to analyze the transcribed interviews and leveraged a combination of video and audio recordings to understand what is involved in searching and selecting apps for smoking cessation.

Results: The process of finding smoking cessation apps is comprehensive: participants explored, evaluated, and searched for information; imagined using functions; compared apps; assessed the trustworthiness of apps and information; and made several decisions while navigating the internet and app stores. During the search, the participants gained knowledge of apps and developed clearer ideas about their wishes and requirements. Confidence and trust in these apps to help quitting remained quite low or even decreased. Although the process was predominantly a positive experience, the whole process took time and energy and caused negative emotions such as frustration and disappointment for some participants. In addition, without the participants realizing it, errors in information processing occurred, which affected the choices they made. All participants chose an app with the explicit intention of using it. After 2 weeks, of the 10 participants, 6 had used the app, of whom only 1 extensively.

Conclusions: Finding an app in the current app stores that contains functions and features expected to help in quitting smoking takes considerable time and energy, can be a negative experience, and is prone to errors in information processing that affect decision-making. Therefore, we advise the further development of decision aids, such as advanced filters, recommender systems and curated health app portals, and make a number of concrete recommendations for the design of such systems.

(*JMIR Hum Factors* 2022;9(2):e32628) doi:[10.2196/32628](https://doi.org/10.2196/32628)

KEYWORDS

mHealth and eHealth; contextual inquiry; decision-making; mobile app search and selection; apps for smoking cessation; mobile apps; mobile phone

Introduction

Background

It is well-established that the toxins in tobacco cause a range of diseases and disorders, often leading to death [1]. The World Health Organization estimates that tobacco kills up to half of its users, which adds up to >8 million people each year [2]. In addition to its major impact on mortality worldwide, tobacco use also results in a great number of morbidities [1]. Smokers die younger, age more quickly, and develop diseases of nonsmokers at a much younger age [3], decreasing the quality of life earlier in life.

Smoking cessation yields specific benefits of reducing fatal and nonfatal vascular, respiratory, and neoplastic (cancer) diseases [4]. Quitting cuts the risk of developing smoking-related diseases, such as lung cancer, by half [5] and increases life expectancy. Regardless of age, quitting smoking is always advantageous to one's health. Smokers who successfully quit smoking before the age of 40 years avoid nearly all the increased mortality risks of continued smoking [4]. After the age of approximately 40 years, every year of smoking prevention saves an average of 3 months of healthy life [6]. Even stopping at the age of 60 years will gain a person 3 years of life expectancy [7].

Mobile apps, which are small software applications that run on mobile appliances, such as smartphones and tablets, are generally regarded as useful tools that aid people in their attempts to quit smoking for several reasons. For example, apps can provide highly individualized and intensive interventions [1,8-11]. Furthermore, apps have the ability to reach large audiences, which makes them cost-effective for both users and suppliers [1,8-11]. Moreover, apps can allow users to tailor interventions according to their personal needs [8]. Finally, apps can reach audiences who might not otherwise seek support [11], in part as apps allow for anonymity [12].

In addition, the persuasive technology literature shows that apps have certain characteristics that make them potentially suitable for supporting behavior change [12,13]. For instance, they can tirelessly continue to try to persuade users without getting annoyed or impatient. They are accessible at any time from any place and consequently able to support people in their behavior change even at night or in the privacy of their homes [1,8-10,12]. Furthermore, people sometimes view their smartphones as digital companions and effortlessly entrust personal information to them [14], thereby facilitating the aiding function of the technology. Finally, apps can present data and graphics, rich audio and video, animations, simulations, or hyperlinked content, enabling users to choose the modality of their preference [12], which could be beneficial for behavior change [15].

Apps for smoking cessation have their own individual characteristics and may vary in terms of usefulness and ease of

use [16,17]; user interface design components, such as navigation, interaction, and appearance [18]; and technical quality [19]. In addition, apps for smoking cessation differ in their approach with regard to the content and its delivery. Hence, there is a fair amount of variation in the main functions of apps [20] and the degree to which apps adhere to clinical guidelines [21] and contain tailoring features [22] or behavior change techniques [16].

Challenges in Searching and Selecting Health Apps

As iOS and Android together account for >99% of the market share in mobile operating systems [23], the Google Play Store and the Apple App Store are, by far, the largest app marketplaces in the field. The total number of apps offered by both these stores is enormous: the Google Play Store offers >3 million apps for potential users, and the Apple App Store has approximately 1.8 million available apps [24]. Although the exact number of available smoking cessation apps is unknown, a person who searches for an app in the Google Play Store, for example, receives the maximum number of search results—250 apps.

Both the Google Play Store and the Apple App Store offer a variety of information cues for each app, such as title, price, rating (stars), ranking (the order in which search results appear in a list), reviews, descriptions, categories, permissions, and the number of installations (only in the Google Play Store). In general, app developers create most of the provided information cues (ie, logo, title, and screenshots). However, the ratings are created by users. In addition, a special type of information in the search results list is the ranking of results. Ranking refers to the order in which the app store presents search results. App store algorithms determine this order, and the exact underlying factors are unclear. However, research suggests that ranking is a reflection of app success, which is, in turn, determined by factors such as the number of languages supported, package size, app release date [25], free app offers, high volume, high user review scores, and continuous quality updates [26]. Although the provided information cues may be informative, tools to guide users through the massive number of results seem to be lacking [27]. At this moment, the visitor cannot use advanced search, filtering, or sorting options in either store. The immense supply of health apps, combined with the lack of tools for refined searching, creates a situation where choosing an app based on anything other than popularity could be considered a challenge.

Related Research

Quantitative studies on uptake, which is the act of downloading and installing smartphone apps in general, have shown that apps with a low price, high ranking, many reviews, and high ratings have the most installations [28] and that high ratings associate more strongly with downloads if customers show a degree of unanimity in their ratings [29]. This implies that these are important information cues for people when choosing apps in general. Diverse qualitative studies have confirmed these

findings. In these studies, participants indicated that they relied the heaviest on ratings, reviews, screenshots, and ranking when choosing various kinds of apps, including apps for smoking cessation [30-32]. Low price or the ability to try an app free of charge are important [32,33], as are the recommendations of others [33,34], preferably given by trusted sources [32].

Specifically, for smoking cessation apps, a few studies have shed light on what people consider important, desirable, or attractive features of smoking cessation apps and which functions people believe to increase engagement. Examples include ease of use, receiving feedback, goal setting, social sharing, competition, and reminders [31,33].

Owing to a recent think-aloud study [35], we now know more about potential users' views on factors such as capability, opportunity, and motivation influencing the uptake of health apps. In this study, Szinay et al [35] found that participants considered searches for health and well-being apps to be difficult, with some calling it a minefield. Furthermore, it was shown that during the search, people pay attention to the look and design, costs, and perceived utility of apps, among others, but primarily to the opinions of others.

These studies provide clear insights into what people generally consider important about apps and which information cues people use before downloading and installing an app. Nevertheless, to the best of our knowledge, it is still unknown how all these insights come together in the process of searching for and selecting health apps in general and apps for smoking cessation in particular. As we know little about the process, we can presently only make assumptions about what the combination of the large supply and lack of tools means for people who want to choose health apps.

Objective

The current gap in the body of knowledge on what people do, experience, and need during the search for mobile apps for smoking cessation creates a need to better understand the process of selecting apps. Understanding the diverse information needs and decision-making processes may ultimately be helpful in creating tools to help people meaningfully select apps. What do people do and experience when searching for an app for smoking cessation? Which information is important to people when choosing an app? How do people use the available information cues in app stores (such as the Google Play Store and the Apple App Store) to obtain the desired information? This study addresses these questions by means of contextual interviews during which people choose an app for smoking cessation. This qualitative approach gives us the opportunity to elicit in situ detailed information to create a rich image based on actual behavior and people's spoken thoughts while in action.

Methods

Study Design

Contextual inquiry is a technique for gathering field data by conducting field interviews with users and studying a task while it is performed in the everyday context. Directly observing the performance of the task enables the revelation of habitual and unconscious practices and is easier for participants as they do

not have to articulate their practices [36,37]. A typical contextual interview, similar to a regular interview, begins with an introduction and some general questions about the participant's situation and then moves on to observation of, and discussion about, the task under study. The researcher not only observes the participant's actions but also pays attention to verbal clues and body language [37]. The distinctive characteristics of a contextual inquiry are the principles of apprenticeship and partnership. In a contextual inquiry, the researcher explicitly assumes the role of *apprentice* and recognizes the respondent as an expert in her or his task. Taking on this role creates a mindset that is focused on curiosity, inquiry, and learning [36]. This mindset is related to working in partnership, which facilitates true collaboration between the interviewer and the respondent to understand the task and motivation of the respondent [36]. This means that the researcher shares thoughts and confusion with the participant on the spot, thus inviting the participant to work together to understand what is happening and why.

Although contextual inquiry originates from and is typically used in contextual design projects [38-40], the method can also be applied to eHealth research [41] on, for instance, mental health [42], healthy eating [43], and persuasive technologies that facilitate healthy lifestyles [44].

Sampling of Participants

We recruited people who wanted to quit smoking, were interested in using an app to do so, and did not currently have or use such an app. Having used a smoking cessation app in the past was not a reason for exclusion. Additional inclusion criteria were (1) owning a mobile device, (2) knowing how to download apps, and (3) being fluent in Dutch.

We recruited participants through posters and social media and by approaching people (who were smoking cigarettes) on the streets in diverse locations in the Netherlands. In addition, we recruited participants through email within our own network. Finally, we used the snowball sampling technique by asking participants at the end of the interview whether they knew someone who might also be interested in participating. To reach our goal of understanding the diverse ways in which people search for smoking cessation apps, we purposively aimed to create variations in age, educational level, and gender.

We created a simple webpage (Qualtrics) in which those interested could leave an email address. Every channel of recruitment contained a link to this webpage. We acquired 20 leads for potential participants whom we sent an information letter. We contacted every lead after a few days to check for interest in participating in the study. Of these 20 individuals, 5 (25%) no longer reacted to our messages, and 5 (25%) had decided not to participate. The reasons stated were not wanting to quit smoking or not wanting to use an app to quit after all, no interest in participating, or practical reasons. Of the 20 individuals, 10 (50%) participated in the study. During analyses of the data, we found that we had reached saturation and, therefore, decided not to recruit additional participants (see the *Strengths and Limitations* section).

Procedure and Data Collection

Some weeks before each interview, we sent participants an information letter, informed them about the use of audio and video recordings, and scheduled an appointment for the interview.

Contextual Inquiry (Interview)

Interviews were conducted face to face (one on one) at a location chosen by the participant. Of the 10 participants, we interviewed 5 (50%) in their homes, 3 (30%) at the university, and 2 (20%) at their workplaces. No one else was present besides the participant and researcher, except for in 1 interview. Researcher SP conducted the interview, and researcher YH, who conducted the other 9 interviews, was present as an observer. Interviews were recorded using a digital voice recorder. During the search for an app, the screens of the participants' appliances were shared with the researcher's laptop (using Mobizen). The footage was recorded using the Microsoft PowerPoint function *Insert Screen Recording*. This captured both footage and sound. The researcher also took notes during the interviews to mainly facilitate revisiting certain remarks and provide a recap at the end of the interview. An interview guide was used to maintain consistency between and direction during the interviews.

Every session started with an introduction explaining the purpose of the study, talking about expectations, asking permission for recording, and answering participants' questions. Participants subsequently provided informed consent on paper.

The introduction was followed by a semistructured interview in which we collected data on age, educational level, and smoking behavior of the participants by asking them. We also talked about prior experiences with eHealth apps, especially for smoking cessation, and about prior experiences with quitting attempts. To get a feel for the motivation of each participant to quit smoking, we used motivation rulers for smoking cessation [45]. On a scale of 0 to 10, we asked participants to indicate the extent to which they considered quitting important, how ready they felt to quit, and how confident they were about quitting. Importance, readiness, and confidence have been associated with smoking behavior change and higher scores, especially on *confidence*, indicating a greater likelihood of attempting to quit [45].

Subsequently, in the contextual interview, we collected data on the process of searching and selecting apps for smoking cessation. We instructed people to search for an app in the way they normally would if we were not present and gave no further instructions on where to start or how to go about the task. We told participants that the task would be completed as soon as one found an app that they considered good, adding that deciding there were no good apps and downloading nothing was also a valid option. We asked the participants to tell us aloud what they were doing, thinking, and feeling. We also asked questions about the task during the search, such as "what is your feeling, when you look at this app?" or "why did you go back to the search results?" (Multimedia Appendix 1 [36,45,46]).

After the participants made their final choice for an app, we jointly created a summary of the entire search process. Doing this together with the participant served as a means of checking

our interpretations. By sharing our interpretations and being honest about interpersonal cues, we aimed to create a valid understanding [47]. In addition, questions we did not ask during the search to not interrupt the participant could be asked here. Before closing off, we informed the participants about the follow-up procedure and planned the date for a follow-up phone interview.

The length of the full sessions (from introduction to completion) varied from 50 minutes to 2 hours and 40 minutes, with an average of 1.5 hours (SD 34 minutes). The duration of the actual searches ranged from 17 minutes to 1 hour and 40 minutes (average 46, SD 26 minutes). It is important to note that this does not necessarily reflect *pure search time*, as, during the search, participants frequently explained their choices and voiced their ideas and thoughts. Therefore, search time is more related to the verbosity of the participant rather than to, for example, the number of apps that were reviewed.

Follow-up Phone Interviews

After 2 weeks from the contextual inquiry, we called the participants over the phone for a final, short semistructured interview. The researcher called the participant at the agreed-upon time. Before starting the interview, we once again asked permission to record the conversation. To do this, we used a digital voice recorder and an Olympus Telephone Pick-up Microphone. Again, we used an interview guide for the topics we wanted to address. Telephone interviews lasted between 10 and 34 (average 19, SD 9) minutes.

In the follow-up interviews, we collected data on the realization of expectations about the chosen app. Some topics we touched upon were as follows: did the participant use the app, and did the app meet the expectations of the participant, given what the participant had learned about the app during the search? In addition, we asked participants whether they had quit smoking (for topics, see Multimedia Appendix 2). Finally, we used the follow-up interview as an opportunity to come back to things participants had said or done during the contextual inquiry, which needed further clarification.

Afterward, all participants received a €15 (US \$16.35) gift voucher via mail as a token of gratitude for their participation.

Data Analysis

Audio recordings of the interviews and the phone interviews were transcribed verbatim using the f4 transcription software. We used Microsoft PowerPoint to create the so-called *process charts* in which we combined corresponding screenshots and participant quotes. These visualizations enabled us to link images on the participants' screens to what people said at that moment (Multimedia Appendix 3). A particular strength of these visualizations was the possibility of seeing that sometimes participants misread or misinterpreted information on their screens. Quantifiable information, such as the number of apps that participants looked at and the scores on the motivation rulers for smoking cessation, was transferred to Microsoft Excel sheets. From there, we translated some data into categories. For instance, we converted information about the number of cigarettes smoked per day into three categories: light smokers

(does not smoke daily), moderate smokers (<20 cigarettes/day), and heavy smokers (≥20 cigarettes/day) [48].

Thematic analysis [49] was used to analyze the transcripts (Textbox 1) and was supported by the use of qualitative data analysis software (Atlas.ti 8 [ATLAS.ti Scientific Software Development GmbH]) and process charts.

Textbox 1. Description of steps in data analysis.

Stage and description

- *Familiarization with the data:* YH transcribed data, read and reread every transcript while listening to the recordings, and created extensive notes and memos on everything that attracted attention. We created, for instance, a memo about the observation that during the search, multiple participants wondered whether certain app features were suitable for them and whether they could see themselves using them.
- *Generating initial codes:* YH marked all possibly relevant text fragments to condense the data and clear out noise. In this step, YH also complemented memos and created new ones. The first transcript was coded independently by both SP and YH. The 2 versions were discussed in detail, and agreement was reached on what and how to code. A final single coded version was created. The remaining transcripts were coded by the first author (YH) while regularly conferring with the second author (SP).
- *Searching for themes:* YH and SP identified the initial main themes, such as starting situation of participants, navigational patterns, and use of information cues to structure the remainder of the analysis process.
- *Reviewing themes:* YH reviewed the initial themes by going through every transcript and process chart, one theme at a time, selecting text snippets and systematically creating headings and ordering fragments under the headings (open coding [50]).
- *Defining and naming themes:* In this step, to refine ideas about the themes and the narrative of the data, YH rearranged the headings, reorganized the text fragments, and reduced the number of headings (axial coding [50]).
- *Producing the report:* YH created the arrangement of the report using the themes on the final classification as headings. The final data analysis was interwoven with the writing process, meaning that we continuously alternated between writing, checking data, adjusting paragraphs, rearranging text, and selecting vivid and appropriate extracts to clarify the report of the results. Multiple iterations of the report were shared, discussed, and refined by all authors. For full, transparent reporting of this study, we used the Standards for Reporting Qualitative Research [51] (Multimedia Appendix 4, [51,52]).

Ethical Approval

Ethical approval for the study was obtained from the institutional review board of the YH's university—the ethics review board of the Tilburg School of Social and Behavioral Science (reference EC-2018.92).

Results

Overview

By analyzing the data from the interviews and contextual inquiries, we identified several facets that play a role in the search for smoking cessation apps (Textbox 2). For the sake of readability and clarity, the report in the *Results* section is structured according to the process steps, and the themes or subthemes are addressed in the description of the process step

they relate to. Furthermore, the *Principal Findings* section contains a descriptive overview and summary of the themes or subthemes.

The remainder of the *Results* section is organized as follows: we start with a description of our participants, their experience with attempts to quit smoking in the past, as well as their previous experiences with smoking cessation aids and eHealth in general. Then, we describe the identified steps of the search process and search thoroughness. Next, we describe the results per process step, focusing on, among others, participants' information needs, actions, decisions, the reasons for those decisions, and participants' search experience. We then describe the transformation of knowledge, wishes and requirements, and confidence in smoking cessation apps throughout the search and across the process steps.

Textbox 2. Facets of searching for a smoking cessation app: themes and subthemes.

Major themes and subthemes

- Search process
 - Extensiveness and thoroughness
 - Decision moments
 - Differences and similarities between process steps
- Information needs
 - Information cue use
 - Functioning of apps
 - Trustworthiness and personal relevance of the information
 - Availability of information
- Information processing and decision-making
 - Activities, cognitive processes, and cognitive load
 - Availability of information
 - Errors in information (processing)
- Transformations
 - Knowledge
 - Wishes and requirements
 - Confidence in apps

Sample Descriptive

The average age of the 10 participants was 41.2 (SD 8.7; range 26-59) years; 6 (60%) were women, and 4 (40%) were men. Of the 10 participants, 4 (40%) had higher education, 4 (40%) had middle education, and 2 (20%) had lower education. Every participant had started smoking as a teenager, at an average age of 16 (SD 1.8; range 13-18) years. This means that the participants had been smoking for 10 to 45 (mean 25, SD 9.3) years. Our sample of 10 participants comprised 4 (40%) heavy smokers, 5 (50%) moderate smokers, and 1 (10%) light smoker. Half of the participants mentioned stress relief as their main reason for smoking. Other reasons were having something to do at certain moments, enjoyment of the taste or the act of smoking, and regarding *being a smoker* as something positive (self-image). Most said that they probably kept smoking as it was a habit and an addiction.

Quitting Smoking

All 10 participants had made serious attempts to quit smoking in the past: 5 (50%) participants made one attempt, 4 (40%) participants made between 2 and 6 attempts, and 1 (10%) participant reported trying 20 times. Some memories of quitting attempts in the past were positive. For example, one of the participants recalled the freedom she felt to be independent of tobacco. Another remembered the fun, game-like aspect of no one noticing that he had quit. However, most recollections of quitting attempts in the past were negative. People remembered how hard it was to quit, how ill-tempered and irritated they felt, and the guilt and shame when the quitting attempt eventually

failed. Some participants specifically mentioned losing faith in their own capability to quit and being afraid of trying again:

I soooooo want to quit smoking, If I had to give it a number it would be a 10, but I am terrified to fail again. [participant 5]

Reasons for wanting to quit again were health (10/10, 100%), the sake of the children (5/10, 50%), and general negative aspects of smoking such as costs, bad smell, and social disapproval. For most participants, the *health reason* was merely a rational, calculated consideration, as most of the participants did not experience any health problems at the time of the interview:

Yes, you see, if I continue smoking the chance of diseases and such is big, so then...But right now I'm fit and healthy. So in the short term that is not a motivation, but in the long run it is. [participant 7]

On the motivation rulers, the participants scored an average of 7.5 (SD 1.35; range 5-10) on the importance of quitting smoking, an average of 6.8 (SD 2.08; range 4-10) on the readiness to quit, and an average of 5.7 (SD 3.55; range 0-10) on being confident that they will quit in the next attempt.

Experience With Smoking Cessation Aids and Apps

Almost every participant had tried some form of smoking cessation aid in the past, ranging from hypnotherapy, acupuncture, and laser therapy to medication, chewing gum, and nicotine patches. Overall, 6 participants had used a smoking cessation app on previous quitting attempts. All participants had fairly low expectations of the benefits of all these aids.

Everyone seemed to feel that quitting is something you need to do by yourself, that it is going to be hard no matter what, and that these aids can be a *helping hand* at most. This sentiment also applied to apps for smoking cessation. Although people found certain functions in smoking cessation apps somewhat useful or motivating, there were more comments on negative aspects, such as the inability of the app to engage them, having to pay to get access to more content, and a lack of interesting functions.

Search and App Selection Process

Overview

The basic steps in the search process were the same for all participants (Figure 1): every search started with entering a search query, which led to a set of results. The next step was to choose a result to obtain detailed information. Subsequently, participants decided to either return to one of the earlier steps or move on to downloading an app. Every participant opened the downloaded apps before deciding to either choose the app or continue the search. All searches ended with participants choosing at least one app they intended to use during their quit attempts.

Although every participant’s search fitted this general process, we also saw some differences. First, we could discern 2 levels of complexity in search flows. Of the 10 participants, 6 showed a simple linear flow. They went from search queries to results, inspected between 2 and 7 different detailed app information screens, and subsequently chose 1 or 2 apps to use. The remaining 4 participants showed a more complex, elaborate flow, with more loops back to the previous process steps, using more search queries, exploring more app information screens, and downloading and discarding more than one app (Table 1).

In addition to the difference in the complexity of the process flow, participants differed from each other in search thoroughness. Some participants (2/10) only scrolled a maximum of 10 apps down in the search results list, whereas other participants (3/10) examined apps in the top 20, and half (5/10) scrolled down even further, sometimes as far as 90 apps down the list. In addition, some participants (7/10) went back to an information screen they had already seen to gain new insights, whereas others (3/10) never revisited app information screens (Table 1).

Figure 1. App selection process flow for smoking cessation apps. Thicker lines indicate more common occurrences.

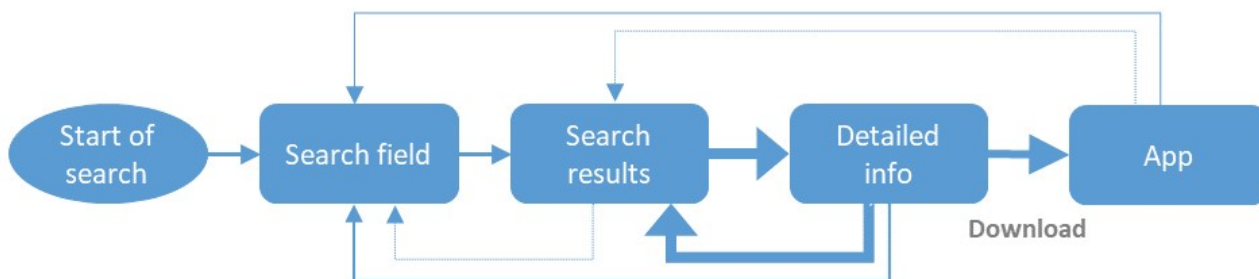


Table 1. Differences in search and app selection process among participants (N=10)^a.

Number	Process flow	Process flow				Search thoroughness	
		Number of different search queries	App information screens (n=85), n (%)	Apps downloaded and opened (n=19), n (%)	Apps chosen for use (n=12), n (%)	Rank of app scrolled to	Revisiting information screens
1	Complex	3	10 (12)	3 (16)	1 (8)	9	Yes
2	Linear	1	2 (2)	1 (5)	1 (8)	12	No
3	Linear	1	6 (7)	1 (5)	1 (8)	12	Yes
4	Linear	1	5 (6)	1 (5)	1 (8)	60	No
5	Complex	4	25 (29)	1 (5)	1 (8)	92	Yes
6	Complex	4	9 (11)	2 (11)	2 (17)	21	Yes
7	Linear	1	7 (8)	2 (11)	2 (17)	28	No
8	Complex	3	11 (13)	6 (32)	1 (8)	16	Yes
9	Linear	1	5 (6)	1 (5)	1 (8)	96	Yes
10	Linear	2	5 (6)	1 (5)	1 (8)	10	Yes

^aApp store could be either the Google Play Store (2, 4, 6, 9, and 10) or the Apple App Store (1, 3, 5, 7, and 8).

Start of Search

The start of the search differed among participants in 2 ways. First, the participants used different devices for the search. Of the 10 participants, 1 started the search on a laptop (switching to a smartphone later on), 2 used a tablet, and the remaining participants searched for an app on a smartphone. Participants starting the search on a laptop or tablet indicated that they thought a bigger screen was somewhat easier for searching and reading.

The second difference was related to the place where the participants started their search. Most participants (8/10) went straight to the app store, whereas some (2/10) participants started their journey in a web browser, using a search engine to visit ≥ 1 website to gain information about apps for smoking cessation before going to an app store:

Because I don't know what I'm looking for, it's nice to spend some time online reading. That gives me some "language", some inkling of what to think of, and after that, I go on to the list of apps. [participant 6]

Search Field

Every participant began by using a search function (either an app store search field or a search engine such as Google). Of the 10 participants, 9 started the search with a Dutch query, and 1 initially used English terms (Multimedia Appendix 5). Seven participants clicked on a query offered by the autosuggestion. Four participants returned to the search function later in the process to use another query (using different terms or switching to English or Dutch) in an attempt to filter the search results (eg, on free or skin) or to search directly for a specific app (by name).

Search Results

After entering a search query, users received a number of results. All participants who searched the Google Play Store and those who searched the Apple App Store with an English search term received only smoking cessation apps in the results. However, Apple users who used a Dutch query received a mix of smoking cessation apps and other unrelated apps, such as Stop Motion apps. One of the participants scrolled down to the 32nd app in the search results list; of the 32 apps, 17 (53%) were not smoking cessation apps.

The information per app in the list of search results in the Google Play Store was limited to a logo, title, rating (number of stars), price (if applicable), and developer name. The Apple App Store additionally provided screenshots but omitted the developer's name. Some participants (4/10) indicated that they thought the information was scarce. For example, they stated that all the ratings were basically the same, and thus unhelpful, and that the other information was hardly useful for making a proper choice. In addition, some remarked that a small text about the functionality of the apps and a means of filtering the results on price were lacking.

The decision people made in this process step was to click or skip an app in the search results list. Most participants (8/10) mentioned that they based their decision on ≥ 1 available

information cue. People most commonly used screenshots (only in the Apple App Store), ratings, price, and the name of apps; however, click-or-skip decisions were also based on logos and developer names. Some participants (2/10) systematically opened app detail pages by first clicking the first app, then the second app, and so on. Hence, these participants did not use any information cues for their click-or-skip decisions.

For the participants who explicitly mentioned using the information cues, we discerned 3 main reasons for clicking or skipping apps. The first was a positive or negative evaluation of some aspects of the app that was reflected directly in the information cues. For example, this was an evaluation of the design of the app based on screenshots, the popularity of the app based on rating, or the trustworthiness of the developer based on developer name (ie, Trimbos Institute). Sometimes, screenshots and app names provided some information about the functionality of an app, on which people based a click-or-skip decision. For example, the term *audiobook* in app names could attract or put off participants. Overall, people clicked an app when the evaluation was positive (design attractive, developer a trustworthy party, rating high, and desirable functionality) and skipped apps when the evaluation was negative (app costs money, design unattractive, rating low compared with other apps, and functionality undesired).

The second reason to click on apps was to check something. For example, several participants clicked on apps to check whether the app was, in fact, a smoking cessation app. In addition, half of the participants indicated some confusion over whether they had already opened detailed information for certain apps at some point during the search. To check this, they clicked or relicked an app in the search results. Furthermore, one of the participants clicked apps out of curiosity and a wish to check what an app was about (triggered by such things as *hypnosis* in the app name, a combination of a trustworthy source and low rating, or a funny name or concept). Finally, one of the participants clicked some apps because of a personal conviction that "one has to check something out to judge it" (participant 10), although the information in the search results did not trigger a particularly positive evaluation of the app.

The third reason for clicking or skipping an app was based on the participant's imagined idea about the working of the app. On the basis of the available information, some participants immediately formed a picture of how the app would work and subsequently clicked on the apps they thought were right for them and skipped the apps they evaluated negatively. In some cases, this interpretation of information led to a decision to skip based on nothing more than a logo, app, or developer name:

The little man here, this one, the green one, kicking his cigarette...I would click that one sooner than this woman with "Quit Buddy". [...] She's going to ask you nicely all the time, I think, or in any case, [she is going to tell you] "well done" all the time. All the time these motivational things. I couldn't take that very well, I think. But that's my first insight. Yeah, I don't know, that [other] one kicks your ass, I guess. [participant 4]

Here: "David Crane, PhD". Somehow that has a weird, nasty...[...] somehow I don't really trust that. Like: here comes [...] mister PhD who will tell us..., He is being paid to promote this. [...] In my head that just turns into something negative. Yeah, that's a personal thing, that I think "what an exaggerated fuss". But, I was like...I have three more to choose from, so I'm just not going to look at this one. [participant 10]

Detailed Information (App Information Screens)

Clicking an app in the list of search results led to a screen with detailed information about a specific app. The information on these screens was far more elaborate than the information in the search results list, containing, among others, screenshots, a description, reviews and additional information about the developer, version, and permissions.

Participants used between 3 and 12 different information cues to gather information about apps. The main sources of information were descriptions, screenshots, reviews, and ratings; however, some participants also considered, for example, the ratio between the number of reviews and ratings, date of the last update, and developer response to reviews. A few participants paid attention to the number of installations (only Google Play Store), and none looked at information about permissions. Some participants showed a clear preference for textual information, others for visual information, and most used both. Most of the time, participants browsed the information; however, sometimes, they went in active search of particular information about, for example, costs or user-friendliness.

While going through the detailed app information screens, participants performed several actions. The most important actions were as follows: (1) they explored information about the functioning of smoking cessation apps, (2) some participants tried to assess the trustworthiness and personal relevance of the information itself, (3) participants formed opinions about diverse functions and characteristics, (4) some participants also imagined what using certain functions would be like for them in practice, and (5) everyone eventually decided to either download an app or leave the detailed app information screen and continue the search. We describe each action in more detail in the following sections.

The primary action on the detailed app information screens was *exploring the information about the functioning of smoking cessation apps* to create a mental image of smoking cessation apps in general and of specific apps in particular. First and foremost, all participants paid attention to what these apps do and how they work by focusing on information about the specific functions of apps, such as time, cigarette, and money counters; challenges; badges; and chat functions. Furthermore, most participants tried to determine whether apps functioned well technically and whether other users were positive or negative about the apps in general and about certain functions in particular. Another important information need was *the price of a free app*. Many participants wanted to know what *hidden costs* were associated with the free apps. These participants were looking for information about the difference between free

and paid versions of the same app; whether one had to start paying over time; and whether paying for an upgrade would get them extra functionality, quality, or just the elimination of annoying advertisements and pop-ups offering upgrades. Finally, several participants looked for information about the quality and professionalism of apps to estimate their trustworthiness. Cues for a trustworthy app could be the name of the developer (known institutions and familiar names generally inspired trust), beautiful design of the app, mention of a scientific foundation, or reactions by the developer to reviews:

There's always a reaction [from developer to reviews] too, right. They always give a...That's definitely positive. Professional. Like, at least he's involved in his own app and taking it seriously. [participant 1]

For some participants, a second action while examining the detailed app information screens was trying to *assess the reliability of information itself*. Half of the participants were engaged in estimating reliability to some degree, which was particularly complicated for reviews:

But then again, I don't really know how that works [...] actually, with apps and with reviews. [...] Yes, [I don't find it credible] that there are so many. [...] I don't really believe that all of that is true, what it says there. Of course, it's also just that it could be someone from Vietnam, who gets paid to write reviews there. I think so. Or, I don't know from which country... [participant 4]

Furthermore, at some point, half of the participants tried to *estimate whether the information was relevant to them* in their search for an app:

"I made a back-up and put it back." [...] Oh, that's just someone who doesn't know how to [...] transfer that to their new phone...That is not applicable to me. She was actually more critical of her inability to install a new phone than of the app itself. [...] That's not a review of the app. Yeah, so then I think, yes, I can sit and read all that nonsense, but it comes down to how it ultimately feels and pleases me in terms of use. [participant 10]

Third, while examining the detailed app information screens, all participants *formulated opinions about functions and characteristics*. These opinions varied from person to person. There was consensus on some functions: the counters and badges were positively regarded by many participants, and the inability to choose one's own quitting date, even if it were in the future, was regarded negatively by most participants. Opinions varied greatly regarding some functions or characteristics:

I'm more one for shock therapy, like: "Stop, stop now! You're getting cancer!", like that. [...] seeing a rotten toe, or something, you know, getting eye cancer from it, that sort of stuff. That impresses me, you know [...] So you have to motivate me, or yeah, punish me, motivate me with my health. [participant 9]

Yeah, you know what the crazy thing is? Yeah, that sounds terrible, I don't know if you've heard it before,

but you know the heart attacks and the lungs, yeah, that doesn't motivate me. Is that bad? [...] This is a very threatening one, with the number of deaths since you stopped smoking and...But that's not my motivation. [...] they've gone out of their way here to make you very afraid in any case, but that doesn't work for me. [participant 5]

Paying for apps was a topic on which all but one of the participants gave their opinion. Of the 10 participants, 4 (40%) were prepared to pay for an app but only if it bought them the extra functionality they desired or if it was a guarantee for a high-quality app; 3 (30%) were not strictly unwilling to pay for an app but thought the free ones would do just fine; and for 2 (20%) participants, paying for smoking cessation apps was an absolute *no go*.

Fourth, next to exploring functions, assessing the reliability and relevance of information, and forming opinions, a number of participants *imagined what using certain functions would be like in practice*. They tried to imagine how and in which situation they would use a specific function:

"Track your cravings and learn how they can get better over time". So apparently, I can register when I'm craving a cigarette. That's kind of interesting because then I can measure it for myself...I know where my weaknesses lie, but [...] I find it interesting because I do think it is fun to do self-examination [...] I do think it's a nice feature, but...I don't think I will make very active use of it, if it's, like, half past one in the morning and I think "I feel like having a cigarette", I don't think I will grab my phone and think, "Half past one in the morning, I'm sitting here on a terrace [...], a glass of wine in my hands and I feel like having a cigarette now. Ohh..." I don't think I'm going to do that. [participant 8]

Finally, at some point in the search, every participant had to *decide to either leave the detailed app information screen or download the app*. This choice was the result of the four abovementioned actions: exploring and imagining resulted in a mental image of smoking cessation apps; opinions about the functions, characteristics, and trustworthiness of the apps; and an assessment of the reliability and personal relevance of information, resulting, in turn, in decisions to either download the app or leave the screen.

In total, the 10 participants opened and left 85 information screens (range 2-25). In some cases, the reason for leaving a detailed app information screen would be practical, such as wanting to see more apps, comparing some apps with others, or an app turning out not to be a smoking cessation app. However, most of the time, people left these screens as the assessment of (some aspects of) the app came out negative. The most common reason for appraising an app negatively was finding a particular function or feature in the app unappealing, unhelpful, or not in accordance with (developing) wishes or requirements. In addition, doubts about the reliability of the app or a certain approach played a role in the negative assessment. Furthermore, bad reviews from others or a small number of reviews and ratings often caused participants to assess an app

negatively and leave the detailed app information screen. For half of the participants, the presentation of information in itself played a role at some point. For the participants, language and spelling errors, poor (automated) translations, and a perceived cluttered structure of text or screenshots were the reasons for leaving a detailed app information screen and continuing the search.

At some point, every participant chose to download an app. The first app was downloaded after participants had viewed, on average, 5 detailed app information screens (range 1-9). Overall, 4 participants downloaded ≥ 1 app (Table 1). Overall, 2 of them (participants 1 and 8) downloaded multiple apps (3 and 6 apps, respectively) to find a specific desired requirement. One of the participants (participant 7) downloaded 2 apps wanting to view them both *live* and then decide which one to keep. One of the participants (participant 6) downloaded an app that he wanted to listen to *just for fun in preparation* in addition to the one he planned to use during the quit attempt. Most participants indicated that they were downloading apps as part of the search process to explore the apps to see what they were like in practice.

The App

All participants opened the apps that they downloaded. Two participants (participants 3 and 6) decided, immediately after opening them, not to explore the chosen apps on the spot. One of them wanted to enter data privately after the interview, and the other did not want to start the trial period at that particular moment. The remaining 8 of participants explored their downloads.

All participants started exploring by clicking on the menu options and buttons to see (and discover) what the app did, how it worked, and what the possibilities were. Exploring the apps resembled the exploration of the information on the detailed app information screens, in the sense that participants stated what they liked or not and what they thought would be helpful. Similarly, the participants imagined whether and how they could potentially use certain functions in practice. In addition to exploring, a number of participants actively sought the functions or features they desired.

Exploring the first download led to the decision to either keep or discard the app. Of the 8 participants, 5 remained (or became more) enthusiastic about their first download after exploration and chose to keep the apps (participants 2, 4, 7, 9, and 10). One participant (participant 5) ran into an *Upgrade to Premium* pop-up, which discouraged proper exploration of the app and made her continue the search without discarding the app. This participant went back to the app store and looked at 17 more detailed app information screens before returning to the initial download and exploring it more thoroughly. After the second exploration, the participant concluded that the app was truly the best one she had encountered and that it actually met her wish or requirement. Of the 8 participants, 2 (participants 1 and 8) discovered something they really disliked about their first app during the exploration and decided to discard the app and continue their search:

I don't even get to choose tomorrow! Or do I have to...? It says here: "Last year." So I can go into the past, but I MUST stay in the now. [...] I don't have a choice. I can't say I want to stop next week because I'm starting medication now for example. [...] They just assume...I want to download the app and they just assume "now you don't smoke anymore". Yes, now I'm already inclined to...I'm curious how that works in the other apps. Whether they also just say "bam"...[...] Well, what irritates me most, or bothers me, is that I am not allowed to choose when I want to stop. [...] I'm just going to find another one.
[participant 1]

From that point on, these 2 participants (participants 1 and 8) changed their way of searching. They had chosen their first downloads as, based on the information they had viewed on the detailed app information screens, they found certain features fun and attractive, could imagine them as helpful, and found the design appealing. After they came across the aspects in their first downloads that were so objectionable (the setting of the quit date in the future and the costs of the app), their search turned into a hunt, really only paying attention to that one requirement. Both decided to keep a downloaded app as soon as they found one that met the requirements.

Choosing to keep an app (and thus stopping the search) was related, in the first place, to satisfaction with certain characteristics and functions but also to a sense of saturation. Half of the participants indicated that they felt they had *explored enough apps*. For some Apple users, this meant that they felt they had viewed the full range of products as the app store returned a limited number of relevant results for a Dutch search query. For a few participants, saturation occurred as their search had taken quite some time, and they had viewed a lot of information. One of the participants was saturated after reviewing a self-pronounced delimited set of the first 10 apps in the search results list.

For a number of participants, in addition to satisfaction with the functions and features and saturation, feeling certain emotions played a role in choosing and discarding an app. Several participants were simply excited enough about the app they had downloaded, opened, and explored to stop searching. One of the participants was surprised to have eventually found exactly what she was looking for. Two participants were tired of searching; 2 others were extremely frustrated during the search and were so relieved when they had finally found something that met their needs that they immediately ended the search:

Whaa! Help. What frustrations...My god. [...] Uhm so no, now I'm like...[...] But what I'll try one more time is to enter "quit smoking" now instead of...See if I get completely different results now. [...] We've already seen this one, we've also seen that one, we've also seen that one...Not this one. [...] [I]m not seeing anything annoying yet, so. I have my health things, I have my milestones. And apparently this is free so then...great. Okay, well, we have an app. And I don't

want to think about it any further now [laughs].
[participant 8]

End of Search

Eventually, every participant ended the search with at least one app and the intention to use it during the next cessation attempt. More than half of the participants felt that they could not still properly judge the app and its usefulness before using it for some time. Several participants indicated that if through use, they would discover that they did not like the chosen app after all, they would have no problem getting rid of the app, switching to another app, or starting to look for other support tools (such as medication or e-cigarettes) for the cessation attempt. This low threshold for discarding the app seemed to be related to the apps being free.

Looking over the process as a whole, across the separate process steps, we observed additional factors that played a role in the choices people made, such as ranking and rating, feelings, and errors in information processing. We describe each factor in more detail in the following sections.

First, the roles of both *ranking* and *rating* in making choices were somewhat ambiguous. Apart from one participant, none literally named *ranking* as important information in their search. Moreover, half of the participants scrolled down further than rank 20 in the search results, and approximately a quarter of the viewed detailed app information screens were those of apps with a ranking >20 (maximum 94). Thus, during the search process, ranking did not seem to play a role for our participants. However, the apps that the participants ultimately chose to use were all in the top 10 in terms of ranking; therefore, ranking did seem to be of influence on the outcome. Similarly, for *rating*, although many participants also viewed information screens of apps with very low ratings (range 2.3-5) and of apps with no rating (because of too few reviews), for some participants, we observed that rating played an important role in the choice of clicking or skipping apps in the search results overview. Moreover, the average rating of the chosen apps was 4.5 (range 3.9-4.8) stars, whereas the average rating of all viewed apps (that had ratings) was 4.3. Once participants arrived on the detailed app information screens, rating seemed less important for some as their focus was drawn to functionalities, design, or other features that excited them or that they considered important.

Second, in addition to rational arguments for choosing to click or skip, leave a detailed information screen, or download or discard an app, almost every participant indicated somewhere in the process that they made a certain decision as something did or did not feel right. For example, one of the participants did not have a good feeling about a particular app while reading the information in the search results and on the detailed information screen. He associated the developer's name with a treatment for alcohol addiction, the app came across as American ("not my favorite...um, people, in terms of attitude and behavior and such" [participant 10]), and he found the use of the word PhD in the developer's name annoying, as well as the mix of Dutch and English in the description. Strictly speaking, none of these things had anything to do with the

content or quality of the app; however, nonetheless, they discouraged him from choosing the app.

Finally, we observed the influence of errors in information and information processing on the decisions participants made throughout the process. For all participants, somewhere in the search process, something went *wrong*. It could be that people missed something in the information, did not read it properly, misinterpreted it, or misremembered it. In addition, the information itself was sometimes unclear, incomplete, or hard to find. As a result, people occasionally drew wrong conclusions and made wrong assumptions. A number of times, we observed that choices (click, skip, download, or discard) were based on a judgment that was formed on information that was misread, misinterpreted, misunderstood, or misremembered or as information could not be found.

In many cases, these kinds of decisions did not necessarily have any kind of impact. For example, one of the participants (participant 3) mixed up all kinds of information she had seen and read. After making the choice, she mentioned that she thought usability was important, as well as the large number of reviews (as to her, that was an indication of many downloads and, thus, popularity, which she considered important). She remembered reading in the reviews of the app she chose that the app was user-friendly. However, the recorded images showed that none of the reviews said anything about user-friendliness. She also remembered that one of the apps she had not chosen had very few reviews. However, the images showed that, of the 6 apps this participant reviewed, the one she referred to was one of the apps that had the most reviews, and the app she had chosen turned out to be one of the apps with the fewest reviews. Thus, it seemed that this participant had misremembered that *negative* features belonged to apps she did not choose, and features she found positive belonged to her chosen app.

In some cases, errors in information (processing) led to a profoundly negative experience or an inferior choice of app. For instance, one of the participants (participant 8) had a very frustrating search caused by not reading carefully and as certain information was hard to find. She mistakenly wrote off several apps that were fully compliant with her requirement for a free app with certain basic functionality. Another participant (participant 10) who did not have a good feeling about a particular app wrongfully assumed the things that made him feel bad about the app (the app was not American but British, eg, and the mix of Dutch and English in the description was caused by an app store functionality and not chosen by the developer). If the participant had not made these errors in information processing and had not written off the app for these reasons, he would have had a higher quality app in this one than the one he ultimately chose.

Thus, over the whole process of ranking and rating, feelings and errors in information processing had some influence on the choices people made. On the other hand, we observed that privacy-related information was not important for any of the participants anywhere in the process. None of the participants viewed the information about permissions on the detailed app information screens. After opening the downloaded apps, almost

all participants instantly agreed with their privacy policies, terms, and conditions. A total of 2 participants first quickly scrolled through the text before giving consent but also immediately indicated the futility of that action:

Yes, actually I always just “agree” [laughs]. I don’t quite feel like reading all the way through, that ehh. [...] I just think, it’ll be fine. [participant 4]

I did read for a while, but then I couldn’t choose anything there. I mean, that was it, so yes, I couldn’t do anything else there except click on it because otherwise I couldn’t continue. [participant 9]

After 2 Weeks

After 2 weeks from choosing an app, of the 10 participants, 6 (60%) had used the app to some extent, of whom 4 (67%) had also quit smoking (Figure 2). Alternatively, one of the participants had quit smoking without using the app. Finally, 3 participants had not used the app and had not quit. Of these 3 participants, 2 had already indicated at the end of the interview that, because of personal circumstances, they were not confident that they would actually start their quit attempt right after the interview (both scored 1 on the confidence ruler), and 1 had not managed to start the quitting attempt, although she was rather enthusiastic about the app and had been moderately motivated to quit during the interview (score of 5 on the confidence ruler).

The 6 participants who had actually used their app had enjoyed occasionally using some functions in the app (the distraction game and the motivation cards) or viewing certain information (the counters and health information). Of the 6 participants, 5 had only used a small number of functions and to a limited extent, and 3 of them indicated that the app could not do much other than count days, cigarettes, and money; however, these participants also immediately admitted that they had not actually explored the app thoroughly. They realized that there might be more functionality available in the apps. For these participants, the app had not played an important role in quitting. However, of the 6 participants, 1 had used the app more extensively and indicated that the app had supported him in his quitting attempt.

In retrospect, what the participants remembered most about finding an app for smoking cessation was that many apps, more or less, offered the same functions and looked similar, making it hard to distinguish among them. We also saw this at times during the search when participants tried to remember the features of a particular app. At such times, it appeared that people mixed up (information about) apps and, in some cases, did not remember whether the app had already been viewed or even downloaded. Combined with hard to find, limited, or absent information, sometimes, the only way to find out about something was to download the app. Consequently, a number of participants indicated that they thought it actually takes (too) much time, effort, and energy (in some cases because of negative emotions) to really look for an app properly. For some, the frustration of the search was still fresh in their minds:

Going back to look for another app? No, no, I found that process so tedious, already after 5 minutes. I’m really not going to do that again, no. Haha, no, I found searching for those apps, oh my god...Those

frustrations all the time, no, oh no, no. [participant 8]

The participants who had used the app intended to leave it on their phones for now, mainly for the counters. Participants who

had not yet used their apps intended to save them for their next quit attempt. Thus, no one expressed any intention of going back to the app store or looking for another app.

Figure 2. App use and quitting success after 2 weeks.

	Quit smoking	Did not quit smoking
Used app	Participants 3, 4, 9, 10	Participants 2, 8
Did not use app	Participant 7	Participants 1, 5, 6

Transformations

From the start of the search until 2 weeks after searching and selecting and, in some cases, using the app, we observed transformations in 3 distinct areas. Over time, participants gained *knowledge* of the workings of smoking cessation apps and simultaneously developed clearer ideas about their personal *wishes and requirements* for an app. However, *confidence and trust* in the ability of these apps to really help while quitting remained quite low or even decreased. We describe the changes in each area in more detail in the following sections.

Before starting the search, every participant could think of at least one or two basic functions (eg, counters and notifications), remembering these from earlier experiences with smoking cessation apps or from stories they had heard from other people. However, none of the participants, including those who used an app in the past, had any knowledge of what currently available smoking cessation apps were able to do and offer. During the search, every participant recognized the basic functions and also discovered new functions and features in smoking cessation apps they had not known or realized existed beforehand. After the search, all participants felt they had a more complete picture of the range of smoking cessation apps, what they can do, and what the landscape looks like.

For all participants, learning about functions and features went hand in hand with forming ideas about what they wanted and did not want from an app. While gaining knowledge, participants developed ideas about what they would like, enjoy, or (on the contrary) find irritating and annoying about an app (wishes), as well as what they thought would or would not help them and, thus, be important in an app (requirements).

The development of wishes and requirements could even continue after choosing an app. The functions and features participants had liked during the 2 weeks of using the apps were often things they had already noticed during the search. However, in some cases, participants were surprised by the fun aspects that they had not seen information about while searching. Notably, a few participants were surprised to find certain functions in the app that motivated them and changed their minds about those functions. For example, one of the

participants gave his opinion about a specific feature while exploring the app during a contextual inquiry:

Well, the way that works, I just find that weird. Because [...] if you accidentally shake your phone, another one of those things will appear. I don't need that. [participant 10]

After 2 weeks, the same participant said the following:

If you have your phone in your hand and you shake it too hard, then it automatically gives those quotes and stuff on the screen, so to speak. Sometimes when you're not even [...] engaged with it, and you pick up your phone, then suddenly there's this thing on the screen, so to speak [...] I think that's a good thing. You don't really get the chance to forget about it, or to let your attention wane, so to speak. So that, yes, for me that works. [participant 10]

The transformation of *confidence* in the helpfulness of smoking cessation apps was slightly more *fuzzy*, with no clear patterns or groups (for an impression of the fuzziness in the changes in confidence, see [Figure 3](#)). Generally, confidence was not very high for any of the participants beforehand. For some participants, this was because of a mediocre experience with these types of apps in the past. For almost all participants, low expectations about the ability of apps to help with quitting seemed linked to low confidence in the ability of cessation aids to help during a quitting attempt in general:

Well, it's certainly not going to be the ultimate remedy. I'm too stubborn for that anyway and I know that I, I have to do it myself. And aside from someone coming and sitting next to me all day and knocking every cigarette out of my hands...There's no way the app is going to do that. [participant 10]

Immediately after choosing an app, participants were asked to estimate their confidence that the chosen app would actually help them quit smoking. For some, confidence had increased slightly compared with the confidence participants indicated having in smoking cessation apps in general before the search; for a few, it was similar; and for one, it had dropped significantly. Although this participant felt that he had chosen

the best app available, he had become disappointed in the landscape of these types of apps through the extensive search:

[...] especially telling that during my search, no really serious things come up. I think that's a very simple fact that says a lot. For example, that there doesn't seem to be an app that costs 100 euros a year. That makes the domain serious, that world, that makes that there is a landscape. That there are things that cost 100 euros, things that are free or a few euros. Then there would be something of a landscape, and now there is not. Actually, we have seen 2½ things. 2½ ways of... That is, a very simple counter and an app that has something of interaction in terms of cravings. [...] Yeah, the disappointment that I feel now at the end, like: "yeah, it's just not there, or something, that [serious] app." [participant 6]

For a couple of participants, the disappointment in the chosen app manifested only after 2 weeks. During the extensive search, they felt that they had looked at enough apps (sometimes *at everything there is*) and had chosen *the best app* from the range on offer, only to discover during use that even the *best* was not very good:

I did hope [that this app could hold my interest]. I think, you know, I was going to do a really good search, and that's what I did with you at the time, but no, [...] no. [...] There's nothing innovative in it. | [...] Maybe I thought, "well, this is it then" because

I chose very consciously [...] and didn't simply take the first one I could find. Then I think, well, this is going [...] to be the Columbus' egg. But it turned out not to be. [participant 9]

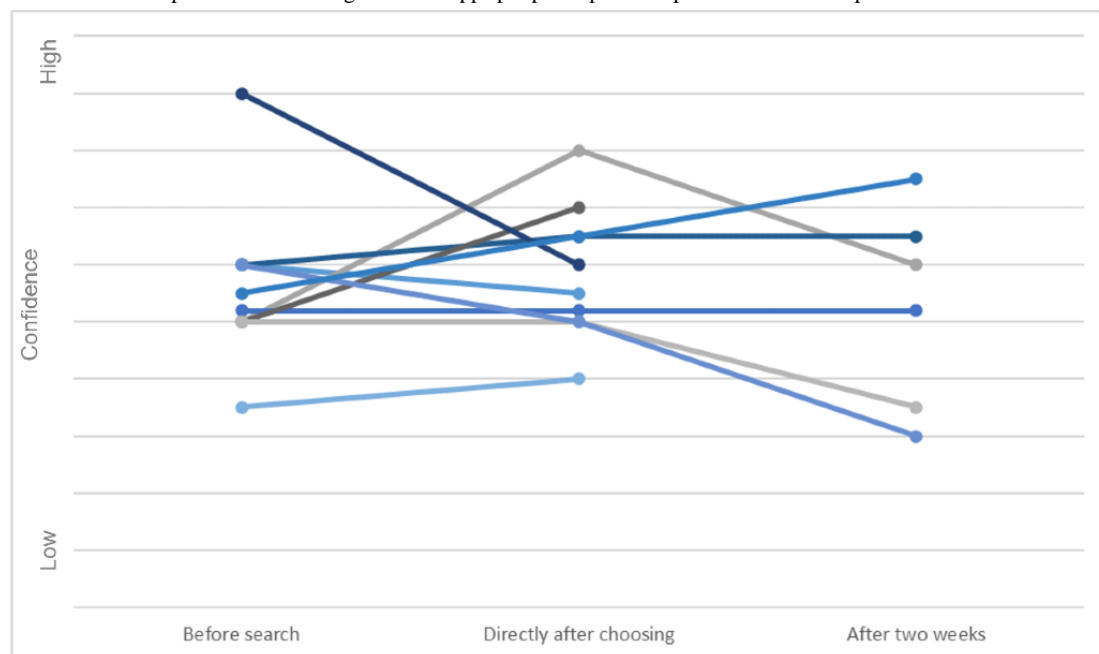
A few participants, who had not had a high opinion of quit-smoking apps 2 weeks earlier, found their lack of confidence confirmed during use:

I don't find the app reliable, because after every day it says: "you will live 60 minutes longer". Then I think: "yeah, bullshit probably". Or just things where you think: yeah, I don't know...this is probably just not true. Anyway, it's kind of funny to see [...] [but] I don't take it very seriously. [participant 8]

For most participants, even for those who were enthusiastic about certain aspects of the app, confidence in the app's ability to help them quit smoking did not increase after 2 weeks compared with the confidence they had immediately after searching. Again, most participants indicated that quitting smoking is simply hard and a matter of perseverance and discipline and that no app in the world can do anything to make it easier:

I do adjust the grade down a bit, to a six [instead of an eight] in the sense that it really helps to stop smoking. [...] you can't stop smoking just by using an app. I mean, there's more to it. But there's nothing the app can do about that. [participant 2]

Figure 3. Confidence in the helpfulness of smoking cessation apps per participant at 3 points in the search process.



Discussion

Principal Findings

This study set out to explore the process of searching and selecting apps for smoking cessation and map the range of actions and the reasons for those actions during the search, focusing on both the information needs and experiences of those who aim to find an app. The empirical findings in this study

have expanded our knowledge of the process, information needs, information processing and decision-making, and transformations that occur when searching and selecting apps for smoking cessation.

With regard to the *process*, we found that participants thoroughly searched for an app that they expected to contribute to smoking cessation. All participants were actively involved in exploring, evaluating, imagining, comparing, searching,

assessing, choosing, and navigating. The comprehensiveness of the search was reflected in several aspects. Many participants continued to look at app information screens and download apps to find something they were somewhat confident in, even if they were fed up or frustrated. The most extensive searches involved using multiple search terms and going back to previously viewed app information screens to discover more and compare apps. Participants viewed many detailed app information screens and scrolled far down the list of search results. No one used a *Take the First* heuristic [30]; one of the participants chose an app after viewing 2 detailed app information screens; however, otherwise, everyone viewed ≥ 5 screens. Many participants read texts thoroughly. Only 1 participant hardly read at all and chose to download apps based on heuristic cues such as ratings and pictures. Most participants also explored the downloaded app as part of the search process. Searches took quite some energy: although there was laughter and participants were generally happy to be looking for an app, for some, the whole process caused negative emotions such as frustration, irritation, and disappointment. Furthermore, several participants indicated fatigue at the end of the search. Overall, it appeared that although confidence in the helpfulness of smoking cessation apps was low, everyone made a real effort to find the *best possible* app.

The search process of our participants was far more extensive than we had expected based on one of the few studies on choosing apps [30]. In that study, only 16% of the participants used a strategy of viewing >1 detailed app information screen before making a choice when choosing (among others) a running app. This former study was conducted in a laboratory setting, used special research devices, and was conducted with participants who did not necessarily have use for a running app. Our participants may have been more invested as they were looking for an app they actually intended to use on their own devices and in their own personal context. It was also notable that all 10 of our respondents chose and downloaded an app with the intent of using it, whereas uptake was found to be far lower in other studies [35]. This result may also be related to the level of investment. Alternatively, it may have been the formulation of the task (*to search for an app like you would do if I were not present*). Although we took care to tell the participants that deciding not to download an app was also an option, emphasizing that choosing an app was certainly not required to end the task, the task may have been leading. Another result we did not expect was the extent to which participants scrolled down the list of search results. It is well-known from research on internet searches that people never scroll down further than the third page [53,54]. However, what is consistent with research in this field is that the first and second results were viewed most often.

Second, with regard to *information needs*, our findings show that participants mainly paid attention to and went in search of information about the functioning of smoking cessation apps. In doing so, they mostly paid attention to what these apps do and how they work, whether apps functioned well technically, whether other users were positive or negative about the apps and their functions, *the price of a free app*, and quality and professionalism of apps. In addition, some participants tried to

assess the trustworthiness and personal relevance of information itself.

Information about the functionalities included in an app (eg, counters, community, Facebook, and coaching) and what features the app has (ie, design and price) was easily found by most participants and could be obtained from descriptions, screenshots, and reviews and by exploring downloaded apps. Information about the content and technical quality of apps could not be gleaned directly from descriptions and screenshots and was, therefore, more difficult to find. Some participants dug deep to assess whether app developers had proper expertise, whether the intervention was good and reliable, and whether it was based on a scientific foundation but often could not find any information about it despite the extensive search. Information about the *true costs* of free apps was equally hard to find.

Nearly everything we have observed in terms of information needs is consistent with previous research. As in previous research, our participants paid attention to features such as monitoring, feedback, goal setting, rewards, reminders and prompts, progress sharing on social media, coping games, health and statistical information, communication style, and ease of use [31-33]. Recently, a study by Szinay et al [35] showed that people also primarily pay attention to these potentially engaging characteristics when searching for health apps. In addition, similar to participants in other studies, our participants also paid attention to immediate look and feel, design, other people's star ratings or reviews of apps (*social proof*), and costs during the search [31,35]. However, the considerable focus on the *hidden costs* of free apps (eg, whether paying for an upgrade would get you extra functionality, quality, or just the elimination of annoying advertisements and pop-ups offering upgrades) is something we have not seen in other studies. This insight is an addition to the factors that people consider important during the uptake of apps for smoking cessation.

Furthermore, this study brings nuances to existing insights from the literature regarding the importance of ranking and rating. Previous research [28,29] has shown that apps with (among other things) high rankings and high ratings are downloaded most often. Although the apps chosen in this study all have high ratings and rankings, this was not what participants paid the most attention to during their search. Participants mainly wanted information about the features of the apps that they expected to be fun or helpful. Individuals looked for specific characteristics of an app (eg, functionality, appearance, and price) and simply started at the top of the search results. Therefore, although ranking was seen by a few participants as a useful source of information for selecting apps, the influence of ranking is clearly noticeable, as starting a search at the top of the list is simply convenient and obvious. This leads to the fact that in the Google Play Store, for example, a few dozen apps at the top of the ranking account for almost half of all downloads [54]. More than 85% of all health apps are found much less often, rarely, or never [55]. This is potentially a loss, as that 85% of apps may contain exactly the functionalities and features that someone is looking for.

Third, in our results regarding *information processing and decision-making*, we observed that participants had to make several decisions during the entire process. In addition to smaller ones, such as choosing search terms, every participant chose to click on or skip apps in the list of search results; leave a detailed information screen or download an app; and finally, after downloading, keep or discard it. To make these decisions, participants needed to understand, interpret, and remember information, form a mental picture of smoking cessation apps, and continually adjust that picture based on new information. Furthermore, choosing an app involved thinking about wishes and requirements and formulating opinions about the functions and features of apps. Some participants also imagined what using certain functions would be like for them in practice. Overall, this seemed to be quite a cognitive load, as without the participants realizing it, they often made mistakes in information processing. A number of times, we observed that choices (click, skip, download, or discard) were based on a judgment formed on information that was misread, misinterpreted, misunderstood, or misremembered, ultimately affecting the final choice for an app. These findings are in line with the known deficiencies in human thinking and decision-making [56,57], including, among others, restricted capacity and forgetting [58]. To the best of our knowledge, the influence of errors in information processing on choosing health apps has previously not been explored and recognized in other studies in this field.

Finally, during searching and selecting, we observed *transformations* in the areas of knowledge, wishes and requirements, and confidence in apps. Knowledge increased from *knowing 1 or 2 basic functions* before starting the search to participants feeling they had a full picture of what smoking cessation apps can do and offer. While gaining knowledge, participants developed ideas about wishes (likes or dislikes) and requirements, which were eventually important in deciding which apps to download, discard, or keep. Notwithstanding this development of knowledge, wishes, and requirements, confidence in smoking cessation apps did not vary much if we compare the participants' estimations before, immediately after, and 2 weeks after the search. For some, confidence was slightly higher immediately after the search, leaving participants rather optimistic. However, that rise was nullified after the 2 weeks of use, with confidence returning to the level before the search or even lower. For some participants, confidence in smoking cessation apps as useful aids had already decreased immediately after the search.

Although it is fully expected that people go through a transformation in knowledge, wishes, and requirements during the decision process [59], to the best of our knowledge, this has not been reported before as an essential part of the search for health apps. We reflect on this in the *Suggestions for Further Research* section. With regard to trust in smoking cessation apps, we confirmed what the study by Regmi et al [60] put forward as a potential threat to smoking cessation apps. In an analysis of strengths, weaknesses, opportunities, and threats, they postulated the loss of trust from users because of the incongruence of perceived app abilities and actual functionalities.

Strengths and Limitations

This study's additions to the literature are primarily the result of our application of contextual inquiry, a method that is not often used in comparable studies. By using contextual inquiries, we were able to study the act of choosing an app in a situation as naturally as possible. Participants could search on their own devices at a place and time that was most convenient for them. This may have increased the likelihood that participants would feel at ease, be honest and open, and understand and accurately remember information, which, in turn, would contribute to data quality [61]. Furthermore, the mindset created by the contextual inquiry's specific basic principles of apprenticeship and partnership facilitated curiosity, humility, interest in, and respect for the respondent, which are generally seen as success factors in conducting interviews [62]. Moreover, close collaboration with participants throughout the research process is thought to produce credible data [63].

We extended our contextual inquiries by video recording the screens of the participants' mobile devices and audio recording their comments. This allowed us to detect that there is often a discrepancy between what people think they see, read, understand, and remember and what actually is on the screens. These double recordings also enabled us to observe that choices are frequently based on the faulty processing of information.

We also consider the inclusion criteria for our participants as a strength of this study. The inclusion of participants in the study who wanted to search for an app to actually stop smoking caused the respondents to be more invested in the task and made the task less artificial. Making observations of actual, realistic behavior in a natural context may have contributed to the ecological validity of the research [64].

Notwithstanding the strengths of the methodology, our chosen approach also has some drawbacks, which create a number of limitations. First, it has been hypothesized that a good rapport between the interviewer and respondent also has downsides and could result in response bias as it causes respondents to ingratiate themselves with interviewers [61]. This could explain why some of our participants indicated that the search during the contextual inquiry was somewhat different from how they would normally search for an app. At the end of the search, 3 participants (participants 3, 4, and 9) indicated that they had chosen more consciously than they normally would have as they had to state aloud why they made certain choices. Three participants (participants 4, 5, and 9) indicated that they had searched a bit more extensively and thoroughly. For 3 participants (participants 2, 7, and 8), the way they had searched for a smoking cessation app was completely different, as normally they would never browse but just go to the app store for a direct search based on an app name. Lastly, 3 participants (participants 1, 6, and 10) searched and found their app in much the same way they would normally do.

In addition, during analyses of the data, we reached saturation [65] when we got to the point where further data collection would not necessarily add anything to the overall story [66]. Even before the analysis of the tenth and last inquiry, we found no new variants of expressions in behavior within the themes or subthemes. The decision to stop further recruitment was

reinforced by the consideration of time investment in recruitment and the chosen methodology. Nevertheless, it is conceivable that studying more people could lead to an even richer description of the search process, people's actions, and their motivations. For example, one of the participants was diagnosed with Pervasive Developmental Disorder-Not Otherwise Specified (PDD-NOS) and surprised us by looking at the information in a completely different way than the other participants.

A further limitation concerns the composition of our sample. As the aim of the study was to better understand the process of choosing a smoking cessation app, recruitment was not about getting a representative sample but about composing a group of people in such a way that we could gain different perspectives. By purposive sampling on factors that could theoretically influence the app choice, such as gender [67-69], age [70], and education level [71], we tried to do just that. However, it is thinkable that different cultural backgrounds or other personal characteristics could provide different, new insights.

Suggestions for Further Research

This study raises several new questions. During the search, participants gained knowledge about smoking cessation apps and developed wishes and requirements. This finding implies that the search process in itself plays a role in the uptake of apps. This raises entirely new questions about the influence of these transformations on the outcome of the search process, selecting an app: how do gaining knowledge and developing wishes and requirements shape the decisions people make, is it an important part of the decision process, does it lead to different outcomes than a search in which no transformations would take place, do these transformations also occur in less active and thorough searches, and what underlying mechanisms are at play? For instance, as all our participants chose an app with the intention of using it, an active and thorough search may have contributed to more satisfaction with the choice [72] and lower uncertainty and thus have increased the intention of using the app [73].

Another potentially interesting question is one regarding the effect of the number of presented search results. The Apple users who used a Dutch search term were presented with significantly less relevant results than those who used an English search term or those who searched the Google Play Store. On average, participants who used Apple explored more app information screens than Android users (mean 11.6 vs mean 5.2) and downloaded more apps (mean 2.6 vs mean 1.2). A number of participants indicated that they liked the fact that there were not so many results but were concurrently puzzled by the limited results and presentation of irrelevant apps. Experimental research with more respondents might explore differences in experiences, feelings, considerations, and decisions among various numbers of search results.

Finally, the matter of initial use is intriguing. Much research in the field has focused on understanding the factors that influence uptake, such as what people find engaging. The goal of many of these studies is to increase uptake by helping users to, for example, obtain information about things that are potentially

engaging [32,33,74,75]. In this study, we saw that participants searched for the functions and features they liked or found useful and that uptake in the sense of *downloads* was high—every participant ended the search with an app and the intention to use it. However, after 2 weeks, we saw that some of the participants had not even opened the app. Despite successful uptake based on expected engaging functions, initial use was thus not guaranteed, let alone actual engagement and continued use. We suggest that it may be worthwhile to investigate what happens between uptake and initial use. It could be useful if further research takes into account the extra step of initial use between uptake and continued use.

Implications for Practice

The results of this study indicate the need to work on the forms of decision support in app stores. We propose a number of suggestions for the design of three obvious solutions to support people in searching and selecting a fitting app for smoking cessation: advanced filters, recommender systems [76], and curated portals [35].

The first solution involves advanced options to filter the search results. In an immense supply, where people want information that is not easy to find, if done properly, filters can make a difference in terms of time, energy, and positive search experience [77]. Choices based on popularity and others' opinions can be made relatively easily by people themselves. Therefore, filters should focus on the content of apps, taking into account the functions and characteristics of the app. With the help of technologies such as natural language processing [78], text analytics [79], and machine learning [80], it is possible to analyze apps in terms of content and identify the characteristics present in the app. Filters and other tools based on the identified characteristics can easily be included in the user interface of an app store, with terms that are relevant, useful, and recognizable to the user, to help the user choose an app that is valuable to them.

The second solution is recommender systems. In this study, we have seen that participants put much effort into figuring out what functions and features they expect will really help them and that they actually find that very difficult to do. Most participants seemed quite unaware of what they needed to support them in their behavior change. Thus, many choices in our study (click or skip, click or download, or keep or discard) ended up being based purely on *a feeling* or on what participants found fun, attractive, or funny. However, there can be a discrepancy between what people indicate to prefer and what actually works well for them [81-83]. The possibility of matching apps and participants with a recommender system could theoretically go beyond matching based on what participants like. Recommender systems could be trained with delayed feedback on the effectiveness of the app on the health behavior change, in this case: smoking cessation. Through this training, the system gradually learns which (functions in which) apps work for whom, optimizing the systems' recommendations on the expected effectiveness of an app, ultimately helping people to find an app that they not only like but that may also work effectively for them.

A solution in the form of curated portals adds value in yet another way when supporting people in choosing an app. First, we have seen that several participants wanted to find a professional, evidence-based app founded on scientific insights. However, information about the quality of apps is almost impossible to find in the app stores. From earlier studies, we know that high-quality apps are scarce to begin with [22,60,84,85] and, therefore, difficult to find in enormous supply. People for whom quality is a criterion would be helped by reliable assistance in choosing. There are reliable sites that users also trust [31,35], such as the GGD App Store in the Netherlands [86]; however, these are not found by users, as this study and previous research have shown [31,35]. An easy-to-find, well-curated site could also help keep people from giving up after a first tried app. It can be a safe and orderly collection where people can return to try a new app if they do not like the first one.

The second argument in favor of curated app portals is data and privacy protection. As in previous studies [35], we also observed that participants hardly glanced at permissions, privacy terms, and conditions. Although people regularly indicate that privacy and data protection are important to them [31,32], in practice, for most, it is not feasible to process and understand this type of information [87]. Even if consumers were to read the incomprehensible terms and conditions, information could be incomplete [88]. Moreover, it has been found that many apps, both free and paid versions, display dismal privacy practices [89,90]. As the use of apps depends on the acceptance of the conditions, and many people are not (or cannot be) aware of the risks [87,88], a reliable, independent party that monitors privacy and data conditions is of great importance.

Conclusions

The empirical findings in this study add insights into the literature on the process, information needs, information processing, and decision-making and transformations in knowledge, wishes and requirements, and confidence and trust that occur when searching and selecting apps for smoking cessation. Currently, finding an app that contains functions and features you expect to help you quit smoking takes considerable time and energy and can even be a negative experience. At present, app stores do not appear tailored to finding suitable smoking cessation apps, and consequently, people who want to quit smoking need to process a lot of information and make a multitude of choices. In this entire process, errors in information processing creep into and affect decisions. Furthermore, although every participant downloaded an app with the intention of using it (uptake), initial use was lower, and subsequent continued use and engagement were even lower. As such, our findings highlight the need for further research into the factors that affect initial use and into the relationship between active, thorough searches and uptake and initial and continued use. Furthermore, our findings stress the importance of developing helpful tools to guide users through the immense supply of health apps, such as advanced filters, recommender systems, and curated health app portals. Among other things, we suggest the creation of filters and recommendations based on app functionalities and curated portals to guide people to high-quality and trustworthy apps. These solutions could make the search process easier, faster, and more enjoyable for people who wish to find an app that is valuable to them and ultimately effective.

Acknowledgments

The authors would like to thank the participants for their time, energy, and candid participation in this study. In addition, they would like to acknowledge the use of DeepL Translator (free version) for translating parts of the text.

Conflicts of Interest

None declared.

Multimedia Appendix 1

Example interview guide contextual inquiry.

[\[PDF File \(Adobe PDF File\), 205 KB - humanfactors_v9i2e32628_app1.pdf\]](#)

Multimedia Appendix 2

Example interview guide follow-up interview.

[\[PDF File \(Adobe PDF File\), 154 KB - humanfactors_v9i2e32628_app2.pdf\]](#)

Multimedia Appendix 3

Example process chart.

[\[PDF File \(Adobe PDF File\), 2889 KB - humanfactors_v9i2e32628_app3.pdf\]](#)

Multimedia Appendix 4

Checklist Standards for Reporting Qualitative Research.

[\[PDF File \(Adobe PDF File\), 142 KB - humanfactors_v9i2e32628_app4.pdf\]](#)

Multimedia Appendix 5

Search queries.

[\[PDF File \(Adobe PDF File\), 120 KB - humanfactors_v9i2e32628_app5.pdf \]](#)**References**

1. Shahab L, West R. Smoking. In: French D, Vedhara K, Kaptein A, Weinman J, editors. *Health Psychology*. Hoboken, New Jersey, United States: John Wiley & Sons; 2010:33-46.
2. Tobacco. World Health Organization. 2020. URL: <https://www.who.int/news-room/fact-sheets/detail/tobacco> [accessed 2020-07-29]
3. Støvring N, Avlund K, Schultz-Larsen K, Schroll M. The cumulative effect of smoking at age 50, 60, and 70 on functional ability at age 75. *Scand J Public Health* 2004;32(4):296-302. [doi: [10.1080/14034940410024176](https://doi.org/10.1080/14034940410024176)] [Medline: [15370770](https://pubmed.ncbi.nlm.nih.gov/15370770/)]
4. Jha P. The hazards of smoking and the benefits of cessation: a critical summation of the epidemiological evidence in high-income countries. *Elife* 2020 Mar 24;9:e49979 [FREE Full text] [doi: [10.7554/eLife.49979](https://doi.org/10.7554/eLife.49979)] [Medline: [32207405](https://pubmed.ncbi.nlm.nih.gov/32207405/)]
5. Peto R, Darby S, Deo H, Silcocks P, Whitley E, Doll R. Smoking, smoking cessation, and lung cancer in the UK since 1950: combination of national statistics with two case-control studies. *Br Med J* 2000 Aug 05;321(7257):323-329 [FREE Full text] [doi: [10.1136/bmj.321.7257.323](https://doi.org/10.1136/bmj.321.7257.323)] [Medline: [10926586](https://pubmed.ncbi.nlm.nih.gov/10926586/)]
6. West R, Shahab L. Smoking cessation interventions. In: Killoran A, Kelly MP, editors. *Evidence-based Public Health: Effectiveness and efficiency*. Oxfordshire, UK: Oxford University Press; 2010:215-232.
7. Tobacco: Health benefits of smoking cessation. World Health Organization. 2020. URL: <https://www.who.int/news-room/q-a-detail/tobacco-health-benefits-of-smoking-cessation> [accessed 2020-08-03]
8. Griffiths F, Lindenmeyer A, Powell J, Lowe P, Thorogood M. Why are health care interventions delivered over the internet? A systematic review of the published literature. *J Med Internet Res* 2006;8(2):e10 [FREE Full text] [doi: [10.2196/jmir.8.2.e10](https://doi.org/10.2196/jmir.8.2.e10)] [Medline: [16867965](https://pubmed.ncbi.nlm.nih.gov/16867965/)]
9. Crutzen R, van der Vaart R, Evers A, Bode C. Public health, behavioural medicine and eHealth technology. In: Gemert-Pijnen L, Kelders SM, Kip H, Sanderman R, editors. *eHealth Research, Theory and Development*. Milton Park, Abingdon, Oxfordshire: Routledge; 2018:111-127.
10. Shahab L, McEwen A. Online support for smoking cessation: a systematic review of the literature. *Addiction* 2009 Nov;104(11):1792-1804. [doi: [10.1111/j.1360-0443.2009.02710.x](https://doi.org/10.1111/j.1360-0443.2009.02710.x)] [Medline: [19832783](https://pubmed.ncbi.nlm.nih.gov/19832783/)]
11. Taylor GM, Dalili MN, Semwal M, Civljak M, Sheikh A, Car J. Internet-based interventions for smoking cessation. *Cochrane Database Syst Rev* 2017 Sep 04;9:CD007078. [doi: [10.1002/14651858.CD007078.pub5](https://doi.org/10.1002/14651858.CD007078.pub5)] [Medline: [28869775](https://pubmed.ncbi.nlm.nih.gov/28869775/)]
12. Fogg BJ. *Persuasive Technology: Using Computers to Change What We Think and Do*. San Francisco: Morgan Kaufmann; 2003.
13. Oyeboode O, Ndulue C, Alhasani M, Orji R. Persuasive mobile apps for health and wellness: a comparative systematic review. In: *Proceedings of the International Conference on Persuasive Technology*. 2020 Presented at: International Conference on Persuasive Technology; April 20-23, 2020; Aalborg, Denmark p. 163-181. [doi: [10.1007/978-3-030-45712-9_13](https://doi.org/10.1007/978-3-030-45712-9_13)]
14. Carolus A, Binder JF, Muench R, Schmidt C, Schneider F, Buglass SL. Smartphones as digital companions: characterizing the relationship between users and their phones. *New Media Society* 2018 Dec 12;21(4):914-938. [doi: [10.1177/1461444818817074](https://doi.org/10.1177/1461444818817074)]
15. Nguyen MH, van Weert JC, Bol N, Loos EF, Tytgat KM, van de Ven AW, et al. Tailoring the mode of information presentation: effects on younger and older adults' attention and recall of online information. *Hum Commun Res* 2016 Oct 21;43(1):102-126. [doi: [10.1111/hcre.12097](https://doi.org/10.1111/hcre.12097)]
16. Ubhi HK, Michie S, Kotz D, van Schayck OC, Selladurai A, West R. Characterising smoking cessation smartphone applications in terms of behaviour change techniques, engagement and ease-of-use features. *Transl Behav Med* 2016 Sep;6(3):410-417 [FREE Full text] [doi: [10.1007/s13142-015-0352-x](https://doi.org/10.1007/s13142-015-0352-x)] [Medline: [27528530](https://pubmed.ncbi.nlm.nih.gov/27528530/)]
17. Oinas-Kukkonen H, Harjumaa M. Persuasive systems design: key issues, process model, and system features. *Commun Assoc Inform Syst* 2009;24. [doi: [10.17705/1CAIS.02428](https://doi.org/10.17705/1CAIS.02428)]
18. Marcus A. *Mobile Persuasion Design: Changing Behaviour by Combining Persuasion Design with Information Design*. New York: Springer; 2015.
19. Thornton L, Quinn C, Birrell L, Guillaumier A, Shaw B, Forbes E, et al. Free smoking cessation mobile apps available in Australia: a quality review and content analysis. *Aust N Z J Public Health* 2017 Dec;41(6):625-630. [doi: [10.1111/1753-6405.12688](https://doi.org/10.1111/1753-6405.12688)] [Medline: [28749591](https://pubmed.ncbi.nlm.nih.gov/28749591/)]
20. Robinson CD, Seaman EL, Grenen E, Montgomery L, Yockey RA, Coa K, et al. A content analysis of smartphone apps for adolescent smoking cessation. *Transl Behav Med* 2020 Feb 03;10(1):302-309. [doi: [10.1093/tbm/iby113](https://doi.org/10.1093/tbm/iby113)] [Medline: [30476293](https://pubmed.ncbi.nlm.nih.gov/30476293/)]
21. Abroms LC, Lee Westmaas J, Bontemps-Jones J, Ramani R, Mellerson J. A content analysis of popular smartphone apps for smoking cessation. *Am J Prev Med* 2013 Dec;45(6):732-736 [FREE Full text] [doi: [10.1016/j.amepre.2013.07.008](https://doi.org/10.1016/j.amepre.2013.07.008)] [Medline: [24237915](https://pubmed.ncbi.nlm.nih.gov/24237915/)]

22. Hoepfner BB, Hoepfner SS, Seaboyer L, Schick MR, Wu GW, Bergman BG, et al. How smart are smartphone apps for smoking cessation? A content analysis. *Nicotine Tob Res* 2016 May;18(5):1025-1031 [[FREE Full text](#)] [doi: [10.1093/ntr/ntv117](https://doi.org/10.1093/ntr/ntv117)] [Medline: [26045249](https://pubmed.ncbi.nlm.nih.gov/26045249/)]
23. Mobile operating system market share worldwide. StatCounter. 2020. URL: <https://gs.statcounter.com/os-market-share/mobile/worldwide/#monthly-201911-202011T> [accessed 2020-12-22]
24. Google Play vs the iOS app store - store stats for mobile apps. 42matters. 2020. URL: <https://42matters.com/stats> [accessed 2020-12-22]
25. Picoto WN, Duarte R, Pinto I. Uncovering top-ranking factors for mobile apps through a multimethod approach. *J Busin Res* 2019 Aug;101:668-674. [doi: [10.1016/j.jbusres.2019.01.038](https://doi.org/10.1016/j.jbusres.2019.01.038)]
26. Lee G, Raghu TS. Determinants of mobile apps' success: evidence from the app store market. *J Manag Inform Syst* 2014 Dec 07;31(2):133-170. [doi: [10.2753/MIS0742-122310206](https://doi.org/10.2753/MIS0742-122310206)]
27. Helf C, Hlavacs H. Apps for life change: critical review and solution directions. *Entertain Comput* 2016 May;14:17-22. [doi: [10.1016/j.entcom.2015.07.001](https://doi.org/10.1016/j.entcom.2015.07.001)]
28. Huang H, Bashir M. Users' adoption of mental health apps: examining the impact of information cues. *JMIR Mhealth Uhealth* 2017 Jun 28;5(6):e83 [[FREE Full text](#)] [doi: [10.2196/mhealth.6827](https://doi.org/10.2196/mhealth.6827)] [Medline: [28659256](https://pubmed.ncbi.nlm.nih.gov/28659256/)]
29. Frick B, Kaimann D. The impact of customer reviews and advertisement efforts on the performance of experience goods in electronic markets. *Appl Econ Lett* 2016 Dec 20;24(17):1237-1240. [doi: [10.1080/13504851.2016.1270399](https://doi.org/10.1080/13504851.2016.1270399)]
30. Dogruel L, Joeckel S, Bowman ND. Choosing the right app: an exploratory perspective on heuristic decision processes for smartphone app selection. *Mobile Media Commun* 2014 Dec 29;3(1):125-144. [doi: [10.1177/2050157914557509](https://doi.org/10.1177/2050157914557509)]
31. Perski O, Blandford A, Ubhi HK, West R, Michie S. Smokers' and drinkers' choice of smartphone applications and expectations of engagement: a think aloud and interview study. *BMC Med Inform Decis Mak* 2017 Feb 28;17(1):25 [[FREE Full text](#)] [doi: [10.1186/s12911-017-0422-8](https://doi.org/10.1186/s12911-017-0422-8)] [Medline: [28241759](https://pubmed.ncbi.nlm.nih.gov/28241759/)]
32. Schueller SM, Neary M, O'Loughlin K, Adkins EC. Discovery of and interest in health apps among those with mental health needs: survey and focus group study. *J Med Internet Res* 2018 Jun 11;20(6):e10141 [[FREE Full text](#)] [doi: [10.2196/10141](https://doi.org/10.2196/10141)] [Medline: [29891468](https://pubmed.ncbi.nlm.nih.gov/29891468/)]
33. Szinay D, Jones A, Chadborn T, Brown J, Naughton F. Influences on the uptake of and engagement with health and well-being smartphone apps: systematic review. *J Med Internet Res* 2020 May 29;22(5):e17572 [[FREE Full text](#)] [doi: [10.2196/17572](https://doi.org/10.2196/17572)] [Medline: [32348255](https://pubmed.ncbi.nlm.nih.gov/32348255/)]
34. Kim H, Kankanhalli A, Lee H. Investigating decision factors in mobile application purchase: a mixed-methods approach. *Inform Manag* 2016 Sep;53(6):727-739. [doi: [10.1016/j.im.2016.02.011](https://doi.org/10.1016/j.im.2016.02.011)]
35. Szinay D, Perski O, Jones A, Chadborn T, Brown J, Naughton F. Influences on the uptake of health and well-being apps and curated app portals: think-aloud and interview study. *JMIR Mhealth Uhealth* 2021 Apr 27;9(4):e27173 [[FREE Full text](#)] [doi: [10.2196/27173](https://doi.org/10.2196/27173)] [Medline: [33904827](https://pubmed.ncbi.nlm.nih.gov/33904827/)]
36. Beyer H, Holtzblatt K. *Contextual Design: Defining Customer-Centered Systems*. Amsterdam: Elsevier; 1998.
37. Holtzblatt K, Beyer H. Field research: data collection and interpretation. In: *Contextual Design: Evolved*. San Rafael, CA: Morgan & Claypool; 2014:11-20.
38. Kaur H, Nori H, Jenkins S, Caruana R, Wallach H, Vaughan JW. Interpreting interpretability: understanding data scientists' use of interpretability tools for machine learning. In: *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems*. 2020 Presented at: CHI '20: CHI Conference on Human Factors in Computing Systems; April 25 - 30, 2020; Honolulu HI USA p. 1-14. [doi: [10.1145/3313831.3376219](https://doi.org/10.1145/3313831.3376219)]
39. Kim K, Kim M, Lee H, Kim J, Moon J. A contextual inquiry of AVEC: power assist wheelchair enhancing communication. In: *Proceedings of the 14th ACM/IEEE International Conference on Human-Robot Interaction (HRI)*. 2019 Presented at: 14th ACM/IEEE International Conference on Human-Robot Interaction (HRI); March 11-14, 2019; Daegu, Korea (South). [doi: [10.1109/hri.2019.8673238](https://doi.org/10.1109/hri.2019.8673238)]
40. Lee CC, Wu FG. Innovative tourism navigation operation process and decision making. In: *Proceedings of the The 2nd International Conference on Software Engineering and Information Management*. 2019 Presented at: ICSIM 2019: The 2nd International Conference on Software Engineering and Information Management; January 10 - 13, 2019; Bali Indonesia p. 259-263 URL: <https://dl.acm.org/doi/10.1145/3305160.3305175> [doi: [10.1145/3305160.3305175](https://doi.org/10.1145/3305160.3305175)]
41. Kip H, Beerlage-de Jong N, Wentzel J. The contextual inquiry. In: Gemert-Pijnen L, Kelders SM, Kip H, Sanderman R, editors. *eHealth Research, Theory and Development*. Oxfordshire, England, UK: Routledge; 2018:167-186.
42. Sakata J, Zhang M, Pu S, Xing J, Versha K. Beam: a mobile application to improve happiness and mental health. In: *Proceedings of the CHI Conference on Human Factors in Computing Systems*. 2014 Presented at: CHI '14: CHI Conference on Human Factors in Computing Systems; April 26 - May 1, 2014; Toronto Ontario Canada p. 221-226. [doi: [10.1145/2559206.2580931](https://doi.org/10.1145/2559206.2580931)]
43. Comber R, Hoonhout J, van Halteren A, Moynihan P, Olivier P. Food practices as situated actionexploring and designing for everyday food practices with households. In: *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. 2013 Presented at: CHI '13: CHI Conference on Human Factors in Computing Systems; April 27 - May 2, 2013; Paris France p. 2457-2466. [doi: [10.1145/2470654.2481340](https://doi.org/10.1145/2470654.2481340)]

44. Xu J, Chen P, Uglow S, Scott A, Montague E. Design challenges and guidelines for persuasive technologies that facilitate healthy lifestyles. *Comput Technol Appl* 2012;3(2):140-147 [FREE Full text] [Medline: 24855501]
45. Boudreaux ED, Sullivan A, Abar B, Bernstein SL, Ginde AA, Camargo CA. Motivation rulers for smoking cessation: a prospective observational examination of construct and predictive validity. *Addict Sci Clin Pract* 2012 Jun 08;7:8 [FREE Full text] [doi: 10.1186/1940-0640-7-8] [Medline: 23186265]
46. Holtzblatt K, Wendell JB, Wood S. The contextual inquiry interview. In: *Rapid Contextual Design*. San Francisco, CA: Morgan Kaufmann; 2005:79-100.
47. Beyer HR, Holtzblatt K. Apprenticing with the customer. *Commun ACM* 1995 May;38(5):45-52. [doi: 10.1145/203356.203365]
48. Gezondheidsenquête. CBS. 2017. URL: <https://www.cbs.nl/nl-nl/nieuws/2017/37/zware-rokers-leven-gemiddeld-13-jaar-korter/gezondheidsenquête> [accessed 2020-12-22]
49. Braun V, Clarke V. Using thematic analysis in psychology. *Qual Res Psychol* 2006 Jan;3(2):77-101. [doi: 10.1191/1478088706qp063oa]
50. Boeije H. *Analysis in Qualitative Research*. Thousand Oaks, CA: Sage Publications; 2009.
51. O'Brien BC, Harris IB, Beckman TJ, Reed DA, Cook DA. Standards for reporting qualitative research: a synthesis of recommendations. *Acad Med* 2014 Sep;89(9):1245-1251 [FREE Full text] [doi: 10.1097/ACM.0000000000000388] [Medline: 24979285]
52. Bunniss S, Kelly D. Research paradigms in medical education research. *Med Educ* 2010 Apr;44(4):358-366. [doi: 10.1111/j.1365-2923.2009.03611.x] [Medline: 20444071]
53. Granka L, Joachims T, Gay G. Eye-tracking analysis of user behavior in WWW search. In: *Proceedings of the 27th Annual International ACM SIGIR Conference on Research and Development in Information Retrieval*. 2004 Presented at: SIGIR04: The 27th ACM/SIGIR International Symposium on Information Retrieval 2004; July 25 - 29, 2004; Sheffield United Kingdom p. 478-479. [doi: 10.1145/1008992.1009079]
54. Lorigo L, Haridasan M, Brynjarsdóttir H, Xia L, Joachims T, Gay G, et al. Eye tracking and online search: lessons learned and challenges ahead. *J Am Soc Inf Sci* 2008 May;59(7):1041-1052. [doi: 10.1002/asi.20794]
55. Aitken M, Clancy B, Nass D. The growing value of digital health: evidence and impact on human health and the healthcare system. IQVIA Institute for Human Data Science, Parsippany, USA. 2017. URL: <https://www.iqvia.com/insights/the-iqvia-institute/reports/the-growing-value-of-digital-health> [accessed 2022-03-04]
56. Tversky A, Kahneman D. Judgment under uncertainty: heuristics and biases. *Science* 1974 Sep 27;185(4157):1124-1131. [doi: 10.1126/science.185.4157.1124] [Medline: 17835457]
57. Kahneman D. *Thinking, Fast and Slow*. New York: Macmillan; 2011.
58. Dörner D, Schaub H. Errors in planning and decision-making and the nature of human information processing. *Appl Psychol* 1994 Oct;43(4):433-453. [doi: 10.1111/j.1464-0597.1994.tb00839.x]
59. Blackwell RD, Miniard PW, Engel JF. *Consumer Behavior (10th)*. Mason, OH: Thomson/South-Western; 2006.
60. Regmi D, Tobutt C, Shaban S. Quality and use of free smoking cessation apps for smartphones. *Int J Technol Assess Health Care* 2018 Jan;34(5):476-480. [doi: 10.1017/S0266462318000521] [Medline: 30226123]
61. Bell K, Fahmy E, Gordon D. Quantitative conversations: the importance of developing rapport in standardised interviewing. *Qual Quant* 2016;50:193-212 [FREE Full text] [doi: 10.1007/s11135-014-0144-2] [Medline: 26792949]
62. Ritchie J, Lewis J, Nicholls C, Ormston R. *Qualitative Research Practice: A Guide for Social Science Students and Researchers*. Thousand Oaks, CA: Sage Publications; 2013.
63. Creswell JW, Miller DL. Determining validity in qualitative inquiry. *Theory Pract* 2000 Aug;39(3):124-130. [doi: 10.1207/s15430421tip3903_2]
64. Lewis-Beck M, Bryman A, Liao T. *The SAGE Encyclopedia of Social Science Research Methods*. Thousand Oaks, CA: Sage Publications; 2004.
65. Saunders B, Sim J, Kingstone T, Baker S, Waterfield J, Bartlam B, et al. Saturation in qualitative research: exploring its conceptualization and operationalization. *Qual Quant* 2018;52(4):1893-1907 [FREE Full text] [doi: 10.1007/s11135-017-0574-8] [Medline: 29937585]
66. Strauss A, Corbin J. *Basics of Qualitative Research: Grounded Theory Procedures and Technique*. 2nd ed. London: Sage Publications; 1998.
67. Lin X, Featherman M, Brooks SL, Hajli N. Exploring gender differences in online consumer purchase decision making: an online product presentation perspective. *Inf Syst Front* 2018 Feb 28;21(5):1187-1201. [doi: 10.1007/s10796-018-9831-1]
68. Bidmon S, Terlutter R. Gender differences in searching for health information on the internet and the virtual patient-physician relationship in Germany: exploratory results on how men and women differ and why. *J Med Internet Res* 2015;17(6):e156 [FREE Full text] [doi: 10.2196/jmir.4127] [Medline: 26099325]
69. Rodgers S, Harris MA. Gender and e-commerce: an exploratory study. *J Advert Res* 2003 Sep 01;43(3):322-329. [doi: 10.2501/jar-43-3-322-329]
70. Peek ST, Luijckx KG, Vrijhoef HJ, Nieboer ME, Aarts S, van der Voort CS, et al. Origins and consequences of technology acquirement by independent-living seniors: towards an integrative model. *BMC Geriatr* 2017 Aug 22;17(1):189 [FREE Full text] [doi: 10.1186/s12877-017-0582-5] [Medline: 28830444]

71. BinDhim NF, McGeechan K, Trevena L. Who uses smoking cessation apps? A feasibility study across three countries via smartphones. *JMIR Mhealth Uhealth* 2014;2(1):e4 [FREE Full text] [doi: [10.2196/mhealth.2841](https://doi.org/10.2196/mhealth.2841)] [Medline: [25098439](https://pubmed.ncbi.nlm.nih.gov/25098439/)]
72. Schwartz B. *The Paradox of Choice*. New York, NY: Harper Collins; 2016.
73. Tudoran AA, Olsen SO, Dopico DC. Satisfaction strength and intention to purchase a new product. *J Consum Behav* 2012 May 23;11(5):391-405. [doi: [10.1002/cb.1384](https://doi.org/10.1002/cb.1384)]
74. König LM, Attig C, Franke T, Renner B. Barriers to and facilitators for using nutrition apps: systematic review and conceptual framework. *JMIR Mhealth Uhealth* 2021 Apr 01:e20037 [FREE Full text] [doi: [10.2196/20037](https://doi.org/10.2196/20037)] [Medline: [34254938](https://pubmed.ncbi.nlm.nih.gov/34254938/)]
75. Russell E, Lloyd-Houldey A, Memon A, Yarker J. Factors influencing uptake and use of a new health information app for young people. *J Technol Hum Serv* 2019 Jan 28;36(4):222-240. [doi: [10.1080/15228835.2018.1536911](https://doi.org/10.1080/15228835.2018.1536911)]
76. Ricci F, Rokach L, Shapira B. Introduction to recommender systems handbook. In: Ricci F, Rokach L, Shapira B, Kantor PB, editors. *Recommender Systems Handbook*. Boston, MA: Springer; 2011:1-35.
77. Bernardi L, Estevez P, Eidis M, Osama E. Recommending accommodation filters with online learning. In: *Proceedings of the ORSUM@ACM RecSys*. 2020 Presented at: ORSUM@ACM RecSys 2020; September 25th, 2020; Virtual Event, Brazil URL: <http://ceur-ws.org/Vol-2715/paper3.pdf>
78. Sarro F, Harman M, Jia Y, Zhang Y. Customer rating reactions can be predicted purely using app features. In: *Proceedings of the IEEE 26th International Requirements Engineering Conference (RE)*. 2018 Presented at: IEEE 26th International Requirements Engineering Conference (RE); Aug. 20-24, 2018; Banff, AB, Canada. [doi: [10.1109/re.2018.00018](https://doi.org/10.1109/re.2018.00018)]
79. Paglialonga A, Schiavo M, Caiani EG. Automated characterization of mobile health apps' features by extracting information from the web: an exploratory study. *Am J Audiol* 2018 Nov 19;27(3S):482-492 [FREE Full text] [doi: [10.1044/2018_AJA-IMIA3-18-0008](https://doi.org/10.1044/2018_AJA-IMIA3-18-0008)] [Medline: [30452752](https://pubmed.ncbi.nlm.nih.gov/30452752/)]
80. Scoecia GL. Automated feature identification for Android apps. In: *Proceedings of the International Conference on Software Engineering and Formal Methods*. 2019 Presented at: International Conference on Software Engineering and Formal Methods; September 18-20, 2019; Oslo, Norway p. 77-84 URL: https://rd.springer.com/chapter/10.1007/978-3-030-57506-9_7
81. Bassili JN. Meta-judgmental versus operative indexes of psychological attributes: the case of measures of attitude strength. *J Personal Soc Psychol* 1996;71(4):637-653. [doi: [10.1037/0022-3514.71.4.637](https://doi.org/10.1037/0022-3514.71.4.637)]
82. Halttu K, Oinas-Kukkonen H. Susceptibility to social influence strategies and persuasive system design: exploring the relationship. *Behav Inform Technol* 2021 Jul 08:1-22. [doi: [10.1080/0144929x.2021.1945685](https://doi.org/10.1080/0144929x.2021.1945685)]
83. Kaptein MC. *Personalized Persuasion in Ambient Intelligence*. Eindhoven: Technische Universiteit Eindhoven; 2012.
84. Haskins BL, Lesperance D, Gibbons P, Boudreaux ED. A systematic review of smartphone applications for smoking cessation. *Transl Behav Med* 2017 Jun;7(2):292-299 [FREE Full text] [doi: [10.1007/s13142-017-0492-2](https://doi.org/10.1007/s13142-017-0492-2)] [Medline: [28527027](https://pubmed.ncbi.nlm.nih.gov/28527027/)]
85. McKay FH, Wright A, Shill J, Stephens H, Uccellini M. Using health and well-being apps for behavior change: a systematic search and rating of apps. *JMIR Mhealth Uhealth* 2019 Jul 04;7(7):e11926 [FREE Full text] [doi: [10.2196/11926](https://doi.org/10.2196/11926)] [Medline: [31274112](https://pubmed.ncbi.nlm.nih.gov/31274112/)]
86. GGD App Store. URL: <https://www.ggdappstore.nl/Appstore/Homepage> [accessed 2021-07-13]
87. Kelley P, Consolvo S, Cranor L, Jung J, Sadeh N, Wetherall D. A conundrum of permissions: installing applications on an android smartphone. In: *Proceedings of the International Conference on Financial Cryptography and Data Security*. 2012 Presented at: International Conference on Financial Cryptography and Data Security; February 27 - March 2, 2012; Kralendijk, Bonaire, Sint Eustatius and Saba p. 68-79. [doi: [10.1007/978-3-642-34638-5_6](https://doi.org/10.1007/978-3-642-34638-5_6)]
88. Huckvale K, Torous J, Larsen ME. Assessment of the data sharing and privacy practices of smartphone apps for depression and smoking cessation. *JAMA Netw Open* 2019 Apr 05;2(4):e192542 [FREE Full text] [doi: [10.1001/jamanetworkopen.2019.2542](https://doi.org/10.1001/jamanetworkopen.2019.2542)] [Medline: [31002321](https://pubmed.ncbi.nlm.nih.gov/31002321/)]
89. Han C, Reyes I, Feal A, Reardon J, Wijesekera P, Vallina-Rodriguez N, et al. The price is (Not) right: comparing privacy in free and paid apps. *Proceedings on Privacy Enhancing Technologies* 2020;3:222-242. [doi: [10.2478/popets-2020-0050](https://doi.org/10.2478/popets-2020-0050)]
90. Polykalas SE, Prezerakos GN. When the mobile app is free, the product is your personal data. *Digit Pol Regul Govern* 2019 Mar 08;21(2):89-101. [doi: [10.1108/dprg-11-2018-0068](https://doi.org/10.1108/dprg-11-2018-0068)]

Edited by A Kushniruk; submitted 23.08.21; peer-reviewed by D Szinay, K O'Loughlin; comments to author 26.10.21; revised version received 24.11.21; accepted 09.01.22; published 14.04.22.

Please cite as:

Hendriks Y, Peek S, Kaptein M, Bongers I

Process and Information Needs When Searching for and Selecting Apps for Smoking Cessation: Qualitative Study Using Contextual Inquiry

JMIR Hum Factors 2022;9(2):e32628

URL: <https://humanfactors.jmir.org/2022/2/e32628>

doi: [10.2196/32628](https://doi.org/10.2196/32628)

PMID: [35436217](https://pubmed.ncbi.nlm.nih.gov/35436217/)

©Ylva Hendriks, Sebastiaan Peek, Maurits Kaptein, Inge Bongers. Originally published in JMIR Human Factors (<https://humanfactors.jmir.org>), 14.04.2022. This is an open-access article distributed under the terms of the Creative Commons Attribution License (<https://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Human Factors, is properly cited. The complete bibliographic information, a link to the original publication on <https://humanfactors.jmir.org>, as well as this copyright and license information must be included.

Original Paper

Factors Influencing Clinician Trust in Predictive Clinical Decision Support Systems for In-Hospital Deterioration: Qualitative Descriptive Study

Jessica M Schwartz^{1,2}, RN, PhD; Maureen George², RN, PhD; Sarah Collins Rossetti^{1,2}, RN, PhD; Patricia C Dykes^{3,4}, RN, PhD; Simon R Minshall⁵, MSc; Eugene Lucas^{1,6}, MD; Kenrick D Cato^{2,7}, RN, PhD

¹Department of Biomedical Informatics, Columbia University, New York, NY, United States

²School of Nursing, Columbia University, New York, NY, United States

³Brigham and Women's Hospital, Boston, MA, United States

⁴Harvard Medical School, Boston, MA, United States

⁵School of Health Information Science, University of Victoria, Victoria, BC, Canada

⁶Weill Cornell Medicine, New York, NY, United States

⁷Department of Emergency Medicine, Columbia University, New York, NY, United States

Corresponding Author:

Jessica M Schwartz, RN, PhD

Department of Biomedical Informatics

Columbia University

622 West 168th Street

PH20 3720

New York, NY, 10032

United States

Phone: 1 212 305 5334

Email: jms2468@cumc.columbia.edu

Abstract

Background: Clinician trust in machine learning–based clinical decision support systems (CDSSs) for predicting in-hospital deterioration (a type of predictive CDSS) is essential for adoption. Evidence shows that clinician trust in predictive CDSSs is influenced by perceived understandability and perceived accuracy.

Objective: The aim of this study was to explore the phenomenon of clinician trust in predictive CDSSs for in-hospital deterioration by confirming and characterizing factors known to influence trust (understandability and accuracy), uncovering and describing other influencing factors, and comparing nurses' and prescribing providers' trust in predictive CDSSs.

Methods: We followed a qualitative descriptive methodology conducting directed deductive and inductive content analysis of interview data. Directed deductive analyses were guided by the human-computer trust conceptual framework. Semistructured interviews were conducted with nurses and prescribing providers (physicians, physician assistants, or nurse practitioners) working with a predictive CDSS at 2 hospitals in Mass General Brigham.

Results: A total of 17 clinicians were interviewed. Concepts from the human-computer trust conceptual framework—*perceived understandability* and *perceived technical competence* (ie, perceived accuracy)—were found to influence clinician trust in predictive CDSSs for in-hospital deterioration. The concordance between clinicians' impressions of patients' clinical status and system predictions influenced clinicians' perceptions of system accuracy. Understandability was influenced by system explanations, both global and local, as well as training. In total, 3 additional themes emerged from the inductive analysis. The first, *perceived actionability*, captured the variation in clinicians' desires for predictive CDSSs to recommend a discrete action. The second, *evidence*, described the importance of both macro- (scientific) and micro- (anecdotal) evidence for fostering trust. The final theme, *equitability*, described fairness in system predictions. The findings were largely similar between nurses and prescribing providers.

Conclusions: Although there is a perceived trade-off between machine learning–based CDSS accuracy and understandability, our findings confirm that both are important for fostering clinician trust in predictive CDSSs for in-hospital deterioration. We found that reliance on the predictive CDSS in the clinical workflow may influence clinicians' requirements for trust. Future

research should explore the impact of reliance, the optimal explanation design for enhancing understandability, and the role of perceived actionability in driving trust.

(*JMIR Hum Factors* 2022;9(2):e33960) doi:[10.2196/33960](https://doi.org/10.2196/33960)

KEYWORDS

clinical decision support systems; machine learning; inpatient; nurses; physicians; qualitative research

Introduction

Background

Clinician adoption of clinical decision support systems (CDSSs) is crucial if best practices are to be integrated into standard clinical workflows. With CDSSs evolving to include machine learning–based CDSSs, the power of machine learning can be leveraged to enhance predictive models of patient risk for a diagnosis or outcome. However, such systems face unique challenges to adoption compared with systems using rule-based logic, which have historically been more widely implemented [1]. A challenge is that the logic behind predictions in machine learning–based CDSSs can be difficult or impossible to make transparent to clinical end users. This has been the focus of much recent research [2-4] in response to the European Union General Data Protection Regulation that effectively mandates a right to explanation of any prediction made based on a person's data [5]. In a study of physicians' ability to understand and explain a machine learning–based CDSS's logic, Diprose et al [6] found that both understandability and explainability were positively associated with *trust*. When the logic behind the predictions was not understood, the physicians did not trust them. Such distrust has been shown to challenge the adoption of machine learning–based CDSSs [7-10], whereas trust is associated with increased intent to adopt machine learning–based CDSSs [11].

Machine learning–based early warning systems, a popular type of CDSS [12], aim to identify patients at risk of deteriorating in the hospital (eg, developing sepsis or experiencing cardiac arrest). These are a type of *predictive CDSS*—machine learning–based systems that make predictions about future patient outcomes or responses to treatment. Predictive CDSSs present more difficulty for clinicians to trust compared with machine learning–based *diagnostic* CDSSs as they require the clinician to trust the accuracy of the prediction even in the absence of objective signs of the outcome. It has been difficult for predictive CDSSs to achieve impactful adoption [10,13,14]. Research indicates that presenting clinicians with a model's overall accuracy is not sufficient for establishing trust in predictive CDSSs [2]. Therefore, how clinicians come to trust and adopt predictive CDSSs remains an area of intense research interest. Moreover, most research on clinician trust has focused on physicians' trust in predictive CDSSs and machine learning–based diagnostic CDSSs [2,6,15]. However, nurses are also target users of predictive CDSSs in the hospital setting [7,16] and may have different perceptions of and requirements for trusting predictive CDSSs.

Others have investigated this topic. For example, a study aimed to explore the factors that influence the integration of predictive CDSSs into clinical workflows and found trust to be an

influencing factor in the emergency department [7]. Others have explored the factors that influence explainability and characterize when explainability increases trust [2]. Another study tested physicians' trust in predictive CDSSs given exposure to different explanations and levels of understanding [6]. The latter two were conducted by referring to simulated as opposed to live implemented predictive CDSSs. Our study is unique from existing research on this topic as it is the first with the objective of elucidating the factors that influence trust referring to an implemented, in-use system in a broad inpatient hospital setting (medical, surgical, and intensive care units).

Objectives

To address this gap in our understanding of how clinicians establish trust in predictive CDSSs and how this might differ by professional group, we explore the experiences of nurses and prescribing providers (physicians, physician assistants [PAs], or nurse practitioners) after the implementation of a predictive CDSS for in-hospital deterioration. Our investigation is guided by a conceptual framework, the human-computer trust framework [17], which accounts for the aforementioned known factors that influence trust—perceived understandability and accuracy. Thus, the objective of our study was to explore the phenomenon of clinician trust among users of a predictive CDSS for in-hospital deterioration by (1) confirming and characterizing the human-computer trust concepts, (2) uncovering and describing any other factors that influence clinician trust, and (3) comparing nurses' and prescribing providers' trust in predictive CDSSs.

Methods

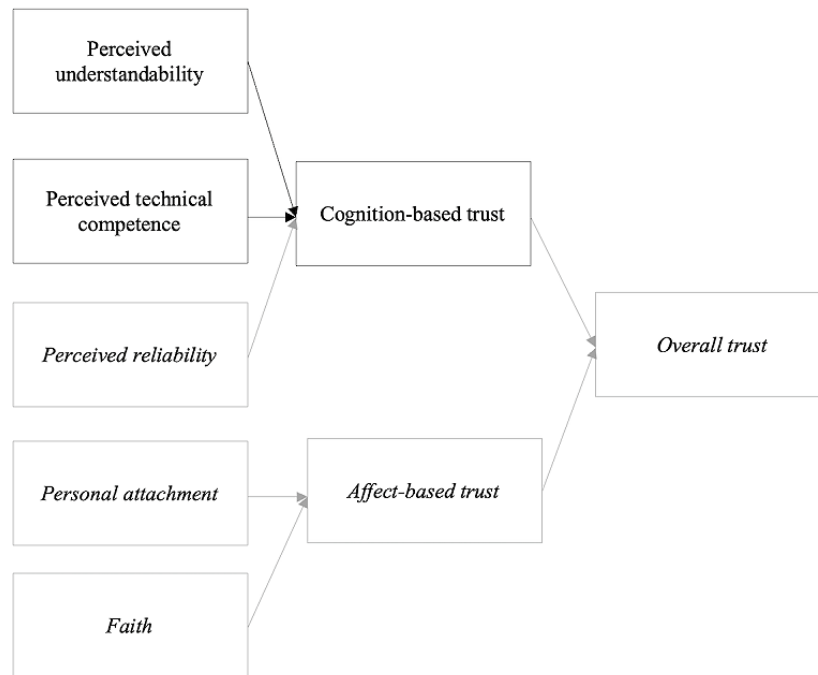
Conceptual Framework

The *human-computer trust* conceptual framework [17] (Figure 1) guided our investigation. In the framework, overall trust is defined as “the extent to which a user is confident in, and willing to act on the basis of, the recommendations, actions, and decisions of an artificially intelligent decision aid” [17]. In this framework, trust is further characterized as cognition-based trust (reflective of the user's intellectual perceptions of the system) and affect-based trust (reflective of the user's emotional perceptions of the system). This study focused on the experience of cognition-based trust and 2 of its concepts: perceived understandability and perceived technical competence. Perceived understandability is defined as “the sense that the human supervisor or observer can form a mental model and predict future system behavior” [17]. Perceived technical competence is defined as “the system is perceived to perform tasks accurately and correctly based on the information that is input” [17]. Although perceived understandability and perceived technical competence are related concepts—ideally, a clinician will

understand a system to judge its accuracy—the inclusion of *perceived* with each concept puts the emphasis on clinicians’

perspectives and how those perspectives influence trust whether they are accurate or not.

Figure 1. Human-computer trust conceptual framework. Perceived understandability, perceived technical competence, and cognition-based trust were investigated in this study.



We did not directly examine perceived reliability because it depends on repeated system use, which was not guaranteed among the participants, and because structural equation modeling led investigators to question its influence on cognition-based trust [17]. Thus, our line of inquiry was focused on participants’ perceptions of their understanding of a predictive CDSS, perceptions of the accuracy of a predictive CDSS, willingness to act based on that predictive CDSS, as well as the factors that influence each of these and their relationship to each other. These concepts of interest were operationalized and explored using a semistructured interview guide (Table S1 in [Multimedia Appendix 1](#)).

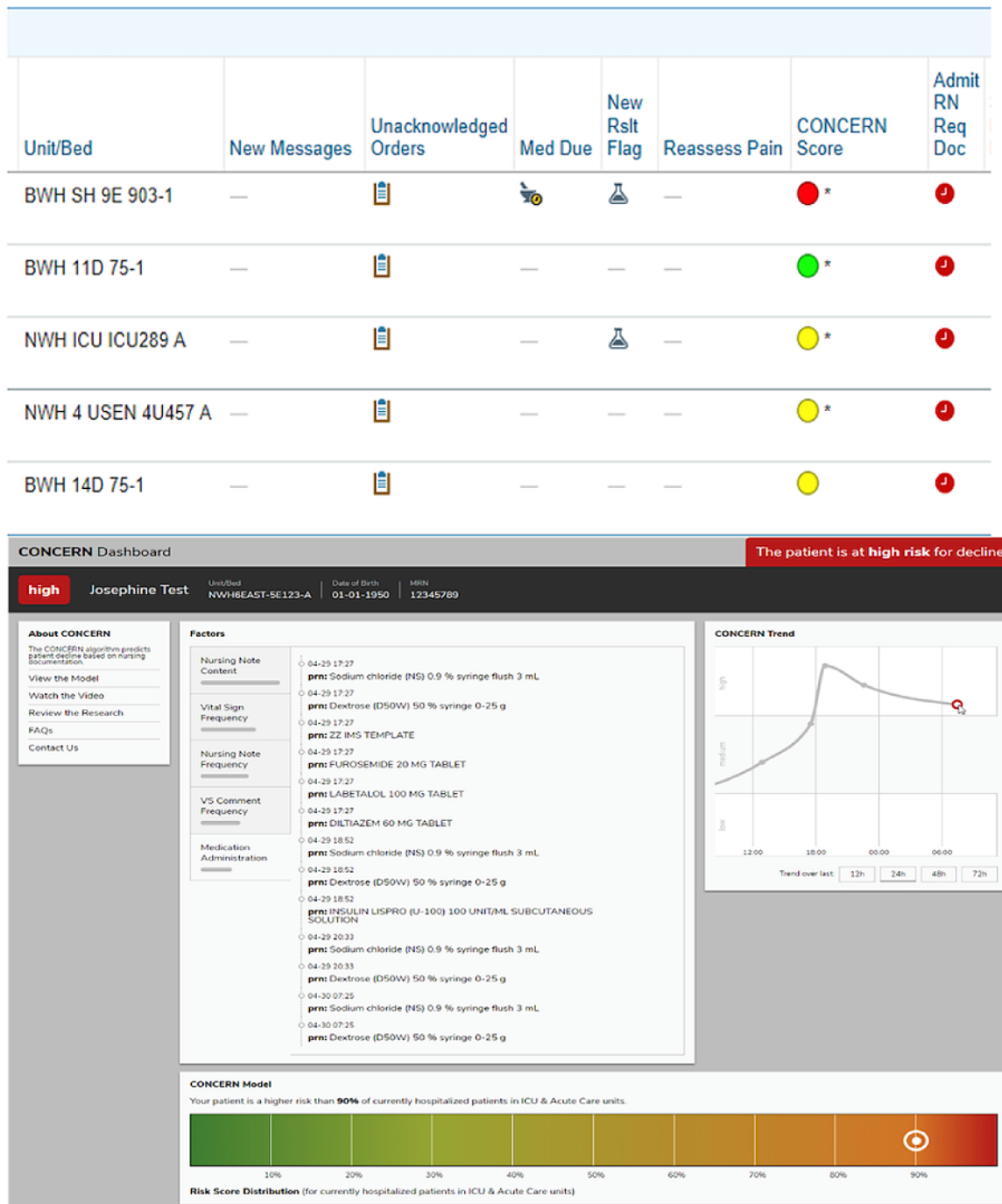
Study Design

Qualitative descriptive methodology guided our methods [18,19], which included directed deductive and inductive content analysis of interview data [20]. A semistructured interview guide (Table S1 in [Multimedia Appendix 1](#)) was developed by the research team and included questions guided by the human-computer trust framework as well as open-ended questions to elicit clinicians’ trust in predictive CDSSs generally and in the Communicating Narrative Concerns Entered by Registered Nurses (CONCERN) CDSS specifically. CONCERN is a predictive CDSS implemented at 2 hospitals within the Mass General Brigham health system that is currently under investigation for its ability to predict in-hospital deterioration (5R01NR016941-05). The system was implemented in July 2020 on 8 pilot units and in September 2020 on 16 additional

study units. The study units included acute and intensive care units, excluding pediatric, neonatal, hospice, emergency, oncology, labor and delivery, behavioral or psychiatric, observational, perioperative, same-day surgery, and plastic surgery units.

CONCERN uses machine learning and natural language processing to model nursing documentation data for predicting patient risk of in-hospital deterioration. As such, it leverages evidence that nurses alter their documentation behavior or use selected language in their narrative notes when they are concerned about a patient’s changing clinical state [21-24]. As shown in [Figure 2](#), CONCERN provides clinicians with a prediction in the form of a colored circle indicating patient risk of deterioration: green indicates low risk, yellow indicates increased risk, and red indicates high risk. By clicking on the color, clinicians open the CONCERN dashboard, which displays the 5 feature (ie, predictor) categories driving the prediction, the relative importance of each in informing that patient’s prediction, the patient-specific documentation contributing to each feature category, a trend line of the patient’s prediction across their admission, where the patient falls along the CONCERN distribution, and links to learn more about CONCERN’s development or provide feedback. The five overarching feature categories used in CONCERN predictions are (1) nursing note content, (2) vital sign frequency, (3) nursing note frequency, (4) vital sign comment frequency, and (5) medication administration.

Figure 2. Communicating Narrative Concerns Entered by Registered Nurses (CONCERN) clinical decision support system.



Participant Recruitment

CONCERN was added to the patient lists of all clinicians working in acute or intensive care units. However, predictions were only displayed for patients admitted to the study units (a random sample of 24 acute or intensive care units). Clinicians in these units received training on CONCERN, and then, 3 months after CONCERN was implemented in all study units (December 2020), clinicians using CONCERN were invited to enroll. Snowball sampling was also used (ie, participants were asked to advertise the study or refer their peers). Clinicians were

not excluded if they had not elected to incorporate CONCERN into their regular practice as we did not want to bias our results toward only those who had a positive perception of CONCERN. Clinicians were characterized as either nurses or prescribing providers (physicians, PAs, and nurse practitioners). The participants received a US \$50 gift card.

Ethics Approval

The Institutional Review Boards at Columbia University Irving Medical Center (AAAR1389) and Brigham and Women’s Hospital (2019P001910) approved this study.

Data Collection

A semistructured interview guide was used to iteratively direct each interview (Table S1 in [Multimedia Appendix 1](#)). Intentionally, we did not introduce the term *machine learning* at the outset of interviews to remain open to clinicians' perceptions of working with predictive CDSSs in practice rather than to potentially bias their responses based on their perceptions of machine learning. Interviews were conducted remotely using Zoom (Zoom Video Communications) outside of clinicians' work hours. Interview length ranged from 20 to 56 minutes (mean length 39, SD 9.5 minutes). Interviews were recorded and transcribed using a Health Insurance Portability and Accountability Act-compliant transcription software (NVivo Transcription; QSR International) and cleaned by the principal investigator (JMS).

We recruited participants until a data saturation table indicated that we had reached data adequacy (ie, no new information was being learned in subsequent interviews [25]) and we had a nearly equal number of nurses and prescribing providers so as to increase confidence in our comparison of findings between the 2 professional groups. Data saturation occurred at the 11th interview (Table S2 in [Multimedia Appendix 1](#)). We continued to recruit up to 17 interviews to balance our sample of professional groups.

Data Analysis

Using both inductive and deductive directed content analysis, we created an initial codebook that defined our codes, established boundaries for their application, and included exemplar quotes [26] (Table S3 in [Multimedia Appendix 1](#)). The primary coder (JMS, a nursing informatics scientist) used this codebook to guide the coding of additional transcripts, revising the codebook as needed. A second coder (SRM, a nonclinical informatics graduate student) completed inductive coding of a purposive sample (half prescribing providers and half nurses) of 47% (8/17) of the transcripts. JMS and SRM met weekly to discuss their findings. Bimonthly debriefings with MG (a qualitative expert with no informatics background) served to achieve a consensus. A third coder (EL, a physician informaticist) completed deductive coding of a purposive sample (half prescribing providers and half nurses) of 35% (6/17) of the transcripts using the concepts of the human-computer trust framework. Intercoder reliability was calculated as Cohen κ

coefficients to measure agreement between the coders performing deductive coding (JMS and EL). Consistent with the qualitative paradigm, the importance of codes was not determined by their absolute frequency [27]. Thus, we report both common and less common perceptions.

Rigor of Data Collection and Analysis

We used multiple strategies for enhancing the trustworthiness of our findings as outlined by Guba [28]. To facilitate credibility (ie, to foster truth in our findings) [28], we used peer debriefing [29] and member checking (verifying emerging codes and categories in interviews with new participants) and assessed our final themes for structural corroboration (to confirm that the findings did not contradict each other). To enhance transferability (ie, the truthful representation of context and sample) [28], we report detailed demographic descriptions of our sample and site and sampled purposively to represent nurses and prescribing providers. To achieve dependability or consistency in the findings [28], we created the codebook and an audit trail documenting all data collection and analytic decisions made throughout the study. Finally, to foster confirmability (ie, reduce bias) [28], the coders practiced reflexivity to identify the researchers' impact on the data. In addition, our interprofessional team of coders and researchers and our purposive sampling strategy that included enrollment of different clinician professions allowed us to triangulate our data; that is, to use multiple perspectives to increase our confidence in the study findings.

Results

Overview

We interviewed a total of 17 clinicians regarding their trust in predictive CDSSs generally and the CONCERN CDSS specifically. Overall, 53% (9/17) of the participants were prescribing providers (8/9, 89% physicians and 1/9, 11% PAs), and 47% (8/17) were nurses. Most clinicians (9/17, 53%) worked on general medicine units or rotations, they had an average of 5.43 years of experience in their current professional role, and an average age of 30.65 years. The participants reported working with the CONCERN CDSS for 1 to 6 months. Most clinicians (14/17, 82%) were recruited from 1 hospital (site A). [Table 1](#) presents the aggregate participant demographics.

Table 1. Participant demographics (N=17).

Demographic variables	Values
Clinician type, n (%)	
Prescribing providers	9 (53)
Physician	8 (47)
Physician assistant	1 (6)
Nurses	8 (47)
Registered nurse	7 (41)
Nurse educator	1 (6)
Current practice setting, n (%)	
Inpatient internal medicine	9 (53)
Cardiology, cardiac surgery, or vascular surgery	4 (24)
COVID-19 (previously internal medicine)	2 (12)
Surgery	1 (6)
Hospitalist	1 (6)
Years in current profession, mean (SD; range)	5.43 (8.59; 0.5-35)
Years at Mass General Brigham, mean (SD; range)	6.12 (7.95; 0.5-32)
Highest professional degree, n (%)	
Medical doctor	8 (47)
Bachelor of Science in Nursing	7 (41)
Master's degree	2 (12)
Age (years), mean (SD; range)	30.65 (8.66; 24-58)
Race, n (%)	
Asian or Asian American	7 (41)
Biracial	1 (6)
White	8 (47)
Not reported	1 (6)
Ethnicity, n (%)	
Brazilian	1 (6)
Chinese	2 (12)
Eastern European	1 (6)
Hispanic	1 (6)
Korean or Korean American	2 (12)
Non-Hispanic	8 (47)
Not reported	2 (12)
Gender, n (%)	
Female	13 (76)
Male	4 (24)
Site, n (%)	
Site A	14 (82)
Site B	3 (18)

Deductive Coding Found Support for the Conceptual Framework

The 2 deductive coders achieved an overall Cohen κ of 0.81. We found support for the 2 concepts of the model: perceived technical competence (Cohen $\kappa=0.77$) and perceived understandability (Cohen $\kappa=0.86$).

Perceived Technical Competence

Clinicians described their trust as being influenced by their perceptions of the accuracy and correctness of CONCERN. For example, a physician said:

The more accurate it is, in my opinion...the more trust I have in the tool. [Physician 2]

Perceived Understandability

Clinicians' ability to understand CONCERN was also confirmed to be an important factor influencing trust. Clinicians described wanting to evaluate the factors contributing to CONCERN to

determine whether they trusted the prediction (also referred to as the "score"):

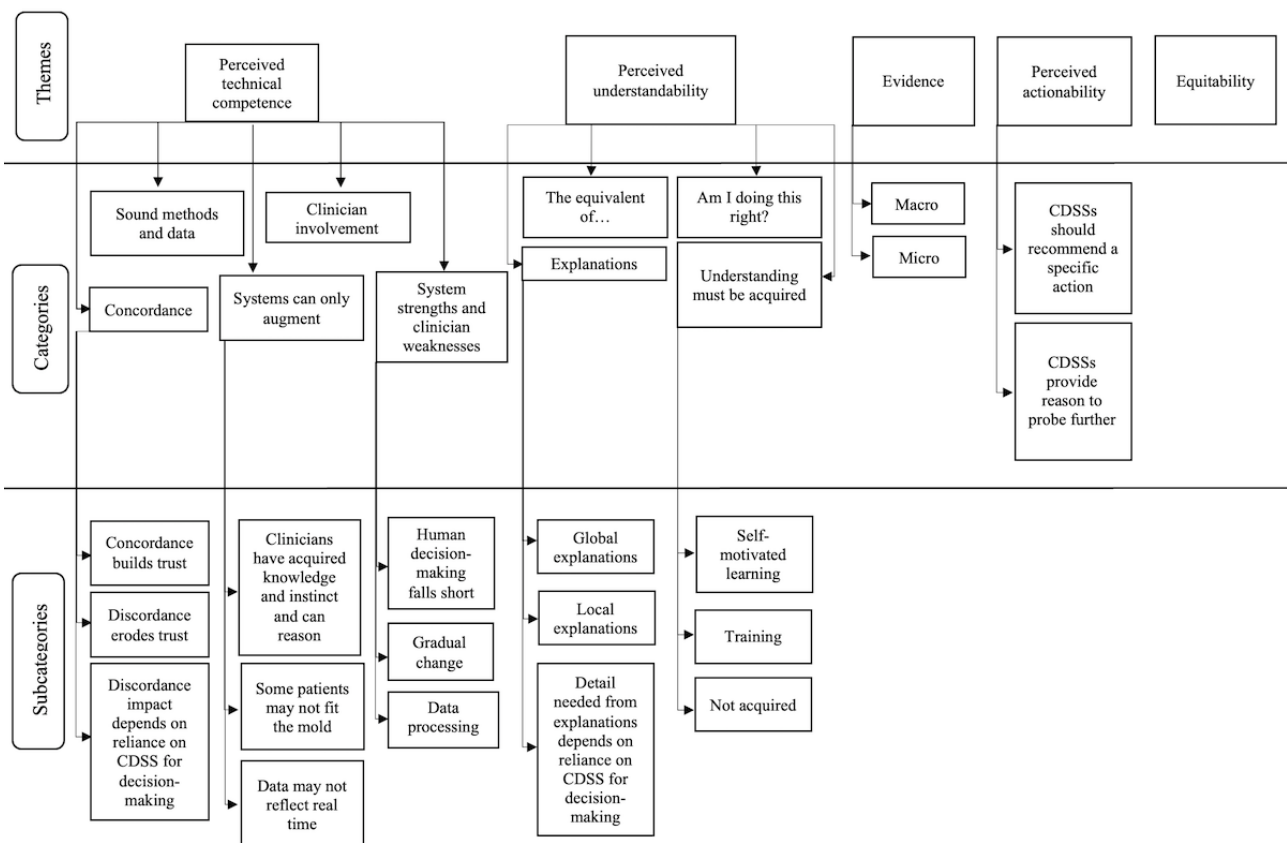
The CONCERN score has changed, like, you know, they're now a yellow or whatever, it might be a good point to be like, oh, what do we think is contributing to that or even reviewing like, because I think there's a way to review, like what, what went into that. And just being like, do we trust this? Do we not? [Physician 4]

Inductive Coding

Overview

The 2 concepts of cognition-based trust, *perceived technical competence* and *perceived understandability*, emerged as themes in the inductive coding. In addition, three new themes reflecting clinicians' trust in predictive CDSSs were identified: (1) *evidence*, (2) *perceived actionability*, and (3) *equitability* (Figure 3). Emergent codes between sites A and B did not differ significantly (Table S2 in Multimedia Appendix 1).

Figure 3. Themes, categories, and subcategories of factors influencing trust. CDSS: clinical decision support system.



Perceived Technical Competence

Regarding *perceived technical competence*, five categories characterized clinicians' perceptions of the accuracy and correctness of CONCERN and predictive CDSSs: (1) *concordance*, (2) *sound methods and data*, (3) *clinician involvement*, (4) *systems can only augment*, and (5) *system strengths and clinician weaknesses*.

Concordance

Concordance between clinicians' impression of the patient's clinical status and CONCERN's prediction emerged as an important factor influencing clinicians' perceptions of the accuracy of CONCERN with (1) *concordance builds trust*, (2) *discordance erodes trust*, and (3) *discordance impact depends on reliance on CDSS for decision-making*.

Concordance Builds Trust

When CONCERN's predictions aligned with clinicians' impressions of the patient, their trust in the system was positively affected:

I felt good that it, very much aligned with how the patient was progressing, whether they were doing well or not doing so well. [registered nurse (RN) 1]

Clinicians also hypothesized that, if there was concordance between their concern for a patient and the CONCERN prediction indicating high risk, they could use the prediction as evidence to support escalating care:

So, I guess in an instance, I could say that, like, oh, this patient's CONCERN score is...red, like, this is just evidence that we need to do intervention. [RN 4]

Discordance Erodes Trust

Conversely, clinicians expressed that a lack of concordance between CONCERN's predictions and their assessments decreased their trust:

I think probably we all kind of take notice of it, but we don't really talk about it because sometimes it doesn't really correlate truly with how a patient is doing clinically. [PA 1]

Similarly, a nurse stated:

That trust could be hindered, say, if I had a patient I was concerned about, and they were a green. [RN 4]

Discordance Impact Depends on Reliance on CDSS for Decision-making

For some clinicians, discordance between the CONCERN prediction and their impression of the patient did not diminish their trust as they viewed CONCERN as just 1 data point that they considered. These clinicians described appreciating the prompt to pay attention to a patient and did not see any harm in an inaccurate prediction:

There have been moments where I'm like, "oh, I've been in there all day, why is it not red?" But again, it's not frequent enough for me to say, "oh, this is garbage." I, I still respect its input [laughs]...it's something I look at at the start of my shift...as the day goes on, I am taking note if there is a change. [RN 3]

The one time it was off, I think, was just there was a lot of documentation happening for other reasons that weren't a good clinical deterioration. There was just a lot of other things going on with this patient that required frequent documentation. And so, it was like a yellow. But again, it was nice to just know, like, oh, I should actually kind of see what's been going on. [Physician 1]

Sound Methods and Data

When referring to the accuracy of CONCERN, many clinicians described their impression of the quality of data used in the model, data that could improve the model, and the modeling methods to varying levels of granularity. A few physicians

expressed their endorsement of CONCERN's methods because it leverages nursing documentation data:

My whole training has kind of just been like, trust your nurses, if they're concerned, you're concerned...And so I think any way that, like, further gives you insight into what is going on, on the nursing perspective, is helpful. [Physician 1]

Others wanted to further scrutinize the rigor of model development:

I'd want to know a bit more about how it was developed, and so let's say the data that CONCERN was trained on was exclusively ICU sepsis and organ failure, mortality, all-cause mortality, let's say...then I would say this tool is only generalizable to the ICU setting, for example. [Physician 3]

Many discussed the data quality, with some clinicians doubting that the frequency of nursing notes predicted deterioration:

[Nursing note] frequency, that only ever happens once a shift...I document my note at the end of my shift whether or not there is a significant event that happened. [RN 4]

Clinician Involvement

Some participants wanted to know that clinicians had been involved in the development of the system or that clinicians would have the opportunity to provide feedback on system performance after its implementation. A nurse said:

I think just with anything, having someone who's actually been there done that is way more, makes it, makes whatever you're developing way more accurate, way more useful, way more diligent. [RN 5]

Systems Can Only Augment

Clinicians identified several limitations to predictive CDSSs relative to clinicians and emphasized that CDSSs were just one of many sources of information that they considered when making clinical decisions. This category is illustrated through three subcategories: (1) *clinicians have acquired knowledge and instinct and can reason*, (2) *some patients may not fit the mold*, and (3) *data may not reflect real time*.

Clinicians Have Acquired Knowledge and Instinct and Can Reason

Clinicians described using their acquired expertise or gut instinct to mitigate patients' risk of deterioration as well as their ability to put the objective data into context:

I feel like a lot of times we just kind of know when somebody is, like, not doing well, especially when we have the same patients often like day to day. [RN 1]

You need those people to look at those numbers that are like patient's tachycardic, heart transplant to say, "yeah, that's abnormal, but it is normal." And in a sense, you can't really computerize that stuff. So that's why a clinician's judgment is so important...you need someone to be thinking, like, what do these numbers actually mean? [RN 4]

Some Patients May Not Fit the Mold

Clinicians expressed skepticism of the system's ability to account for unique or complex patient characteristics:

Because it's so complicated, it's quite, I mean, I hesitate to say unique, but there are a lot of a lot of factors in place. And it would be hard for a training data set to include enough patients who were similarly complex for it to have, let's say, three hundred patients with infection in the right rib and the left shoulder and the left knee all at once. And for it [the system] to kind of know what to do at that point. [Physician 3]

Data May Not Reflect Real Time

Clinicians frequently described scenarios in which data that the system uses would be missing. Examples included emergencies, rapid deterioration, new patients, or when clinicians are burdened by work or documentation:

I'm not saying that systems like this aren't smart, but I just feel like so much of it depends on what's going on in that moment. And a lot of times, you know, our documentation isn't always like right up to date with what's going on at the moment. [RN 7]

System Strengths and Clinician Weaknesses

Clinicians also identified several reasons why predictive CDSSs might be more accurate than an expert clinician in predicting risk of in-hospital deterioration. This category is illustrated through three subcategories: (1) *human decision-making falls short*, (2) *gradual change*, and (3) *data processing*.

Human Decision-making Falls Short

Clinicians described their own and their colleagues' limited abilities to make accurate predictions when they are tired, when they have limited experience (eg, either years in practice or with a particular patient population), when they are overly burdened, or when they are not "good" clinicians. For example, many clinicians mentioned that they would probably rely more on predictive CDSSs during the night shift, when they are assigned to care for more patients:

I'm like doing night coverage, so I don't know the patients as well, so maybe I would, in that setting, be more reliant on a tool like that. [Physician 4]

I'd say [CDSSs would be better than a clinician when] someone, a novice in their role. Like July [laughing about when new residents begin] or any new nurse or anything like that. [RN 3]

Gradual Change

Clinicians described predictive CDSSs as better equipped to make predictions in situations where the change to the patient's state is gradual rather than rapid:

Maybe the algorithm's better at like kind of like nudging us to just like readdress some things that maybe are changing minutely day to day, so we may miss if we're if we're not, like, really aware of the trend. [Physician 4]

Data Processing

Clinicians also recognized that there are large volumes of data to synthesize in the electronic health record (EHR) and that systems may be better equipped to process that volume of data, especially from clinical notes. Exemplar quotes are presented in Table S4 in [Multimedia Appendix 1](#).

Perceived Understandability

Perceived understandability was characterized by four categories: (1) *explanations*, (2) *understanding must be acquired*, (3) *the equivalent of...* (physicians only), and (4) *Am I doing this right?* (nurses only).

Explanations

Much of the clinicians' discussion of understanding CONCERN specifically and predictive CDSSs generally involved explanations of system logic and individual predictions. This category is presented in the form of three subcategories: (1) *global explanations*, (2) *local explanations*, and (3) *detail needed from explanations depends on reliance on CDSS for decision-making*.

Global Explanations

Clinicians wanted global explanations, meaning information on how the CONCERN model calculates predictions generally:

So my questions are like, well, what kinds of phrases and words and how often, you know, is the is the CONCERN tool looking back? Is it, are they looking at one note? Are they looking at three notes? And when you say vital sign frequency, what does that mean? [Physician 7]

I think [to understand an algorithm like CONCERN] just more what it takes into account, whether it's you know, their vital signs or their lab values, I don't really know how it calculates, if they're flagged as yellow or green. [RN 7]

Local Explanations

Clinicians also wanted explanations for individual patient predictions provided at the point of care. A physician said they want to see "the vital signs or the whatever that is making the score change" (Physician 3). A nurse said they would look for "what piece of it is causing the algorithm to say that the person's not stable" (RN 2).

Detail Needed From Explanations Depends on Reliance on CDSS for Decision-making

This third subcategory emerged from some clinicians stating that they did not need detailed explanations of CONCERN as it was just 1 component of their assessments:

The fact that it's, it's an extra data point that's available to me there doesn't make me so concerned about, well, you know, how does the machine learning work and when what goes into this? To me, I'm like, well, I understand what machine learning is and I understand that it helps me better inform some of my clinical decisions and maybe gives me like an extra reason to, to double check my, my own clinical

assessment. So, in that sense, like I feel like it's been a sufficient enough information for me. [Physician 8]

Understanding Must Be Acquired

Clinicians described the various ways in which they came to understand CONCERN or did not understand it. This is described in three subcategories: (1) *self-motivated learning*, (2) *training*, and (3) *not acquired*.

Self-motivated Learning

In this subcategory, some clinicians described themselves as being self-motivated to learn. For example, they may have seen a poster or received an email about CONCERN that prompted them to read about it, look at the predictions more frequently, or investigate the dashboard. Exemplar quotes are presented in Table S4 in [Multimedia Appendix 1](#).

Training

Some clinicians previously participated in CONCERN design focus groups, which they described as a helpful form of training, whereas others received formal training. Clinicians felt that formal training should be part of onboarding new staff. Exemplar quotes are presented in Table S4 in [Multimedia Appendix 1](#).

Not Acquired

Conversely, some clinicians had a poor understanding of CONCERN, with a few who felt that they did not receive education expressing frustration about this. Exemplar quotes are presented in Table S4 in [Multimedia Appendix 1](#).

Although all previous categories were informed by both prescribing providers and nurses, 2 categories were profession-specific: only physicians used analogies, and only nurses were concerned about how their documentation affected predictions.

The Equivalent of...

In this category, some physicians used analogies when describing their understanding of CONCERN:

I literally think of it the same as imaging, like I usually rely on radiology reports, because I'm not a radiologist, but I do like to look at the images myself because...you can sometimes have a different context for what you're looking for that the radiologist doesn't know. [Physician 4]

Am I Doing This Right?

Knowing that their documentation would inform the CONCERN prediction, some nurses wanted to know that they were not missing something that would make the CONCERN prediction more accurate. Some said that they had or would change their documentation behavior in attempts to make the prediction reflect their impression of the patient:

I feel like I do try to put stuff in that's like part of the CONCERN score, but it doesn't always, like the CONCERN score doesn't always reflect it, so then I'm like, I'm not sure I'm putting in the data correctly? Or like I'm just not putting it in the right comment boxes or like filling out my notes, you know,

I don't know if, like, I'm the one who's not raising that level of concern because I'm just not putting, I'm not, like, doing the algorithm correctly where it would recognize it as a concern. [RN 6]

Finally, new themes emerged from the data analysis that do not map onto the conceptual framework. These included (1) *evidence*, (2) *perceived actionability*, and (3) *equitability*.

Evidence

The *evidence* theme emerged from clinicians' discussions of how evidence of CDSSs positively affecting patient care would increase their trust in the system's predictions. In all, 2 categories emerged: *macro* and *micro*.

Macro

In *macro*, scientific evidence of the impact of a predictive CDSS on patient care was important for facilitating trust:

I think, really like a study showing that the score has been used and the evidence behind it...if it's published and peer reviewed, I think I definitely, personally I'd be more more inclined to use it. [Physician 6]

Micro

Clinicians also described the importance of anecdotal reports of positive impact:

I think anecdotally, like if others I know said, "hey, you know, I happened to catch this patient who was deteriorating and we were actually able to like, you know, get involved early and we were able to prevent this patient from either a rapid or like likely ICU transfer." I think those things pull like a lot of weight. [Physician 1]

Perceived Actionability

Some clinicians wanted a clear recommendation for an action to take to trust the predictive CDSS (which CONCERN does not provide). Others discussed how CDSSs provide reason for them to further examine a patient's clinical status.

CDSSs Should Recommend a Specific Action

Some clinicians expressed a desire to know what to do with the patient's risk score to trust it:

Understanding how the predicting part comes in, I think would give me more confidence...some sort of like if/then tool, so if the score is greater than this, then you should take this kind of action. [Physician 3]

CDSSs Provide Reason to Probe Further

Clinicians also noted that a CDSS prediction indicating elevated risk had prompted or would prompt them to investigate further, either via EHR data review or discussion with another team member:

If you're talking about trust, I feel like me looking at a red patient...it would boil down to, OK, this patient's red, I want to look into their chart. [RN 4]

If I saw a red or anything other than green, I'd click on that patient, look at their flowsheets and then if

they were like tachycardic, two hours ago they were not, then I would go and actually visit the patient and check with the nursing team to see if they had any concerns. [Physician 3]

Equitability

A clinician expressed the importance of the predictive model being equitable:

The one caveat [to machine learning] is it could, if it uses, you don't know exactly what data it uses, and I would be interested in studies that explore whether that machine is systemically racist or classist or whatever...So, some sort of study to make sure it's equitable to all populations is important. [Physician 6]

Discussion

Principal Findings in the Context of What Is Known

Our qualitative descriptive investigation using the human-computer trust framework [17] produced broad and deep characterizations of nurses' and prescribing providers' trust—and distrust—in predictive CDSSs. We confirmed that perceived understandability and perceived technical competence influence clinicians' trust in predictive CDSSs as well as identified additional factors: evidence, perceived actionability, and equitability. Furthermore, we found profession-specific factors characterizing the relationship between understandability and trust.

Although we focused on cognition-based trust, our findings have implications for reconceptualizing the human-computer trust framework. In each interview, the clinicians were asked what would increase or decrease their trust in CONCERN. In all, 3 concepts of the framework (perceived reliability, faith, and personal attachment) were not identified by the participants. However, these concepts might be more dependent on sustained system use, which not all of the participants had. Other works conceptualize trust as being influenced by an individual's propensity to be trusting [11]. Although this concept also did not emerge definitively in our study, it is possible that the clinicians who described being self-motivated in their learning about CONCERN were indirectly demonstrating a propensity to be trusting.

Much has been written about the importance and perceived trade-off between accuracy and understandability in machine learning-based CDSSs [30,31]. Our investigation found that both are important and provided context for clinicians' desires for each in the case of predictive CDSSs. As CONCERN was implemented during the COVID-19 pandemic, there were limited opportunities for in-person education, and there were increased demands on the clinical staff. This may have contributed to clinicians having a poor understanding of CONCERN and its global and local explanations on the dashboard. In fact, when some clinicians were asked how they would determine that the CONCERN tool was accurate, they answered by expressing a desire to understand it more thoroughly—indicating a primacy of understandability over accuracy, as others have found [2]. This preference may differ

for machine learning-based diagnostic CDSSs, as hypothesized by Diprose et al [6].

However, delivering an accurate and *desirable* explanation of machine learning logic remains a challenge. When describing their desire for local explanations, many clinicians indicated an orientation toward rule-based causal logic. They wanted to know the one feature or value that made the patient's prediction yellow or red. In the case of many predictive CDSSs, such simplifications are not possible, and an interpretation of causation would not be accurate. When our team iterates on the CONCERN design, we will look to explanation design frameworks such as that outlined by Barda et al [32] to optimize the impact of explanations on understandability. However, long-term strategies aimed at increasing the education that clinicians receive on machine learning are also likely needed, as others have also reported from their investigations [7,33,34].

We found that some of the factors influencing clinicians' perceptions of system accuracy (ie, perceived technical competence) differed from findings in previous research. Tonekaboni et al [2] reported that clinicians would like to see a certainty score presented with the CDSS prediction; however, no clinician requested this or any type of accuracy metric in our interviews. When prompted, they said that an accuracy metric would be helpful, but differences may be attributable to context. Tonekaboni et al [2] interviewed clinicians referring to simulated rather than implemented predictive CDSSs. We found that clinicians primarily judged the accuracy of CONCERN against their own impressions of their patients' risk of deterioration, which may be what clinicians do in real clinical care. In fact, this was suggested by clinicians in the study by Tonekaboni et al [2].

Importantly, many of the categories that inductively emerged in this study align with others' findings. For example, Sandhu et al [7] reported that “even when physicians did not trust a model output, they still reported paying closer attention to a patient's clinical progression,” aligning with our category *CDSSs provide reason to probe further*. Elish [35] also found that evidence was important to clinicians, particularly “anecdotal evidence and discussions of specific cases and patient outcomes,” aligning with our category *micro evidence*. In addition, many have highlighted the importance of engaging clinical end users throughout development and implementation [7,35-38]. However, most predictive CDSS studies do not report involving clinicians in development, indicating that this is an area for future work [39].

Others have warned about an overreliance on inaccurate machine learning-based CDSS predictions or classifications [30,40]. In fact, Jacobs et al [15] found that clinicians trusted incorrect recommendations. Similarly, Cabitza et al [40] argued that clinical users of machine learning-based CDSSs using EHR data need to be aware that data “quality is far from optimal” and warned clinicians about losing awareness of important clinical factors not present in the EHR. However, the clinicians in our study did not show a propensity to overrely on the CONCERN predictions and indicated that they recognized predictive CDSSs' shortcomings.

Only nurses in our study wanted to understand how to document “correctly” for the CONCERN score, with some indicating that they would change or had already changed their documentation behavior to make the CONCERN score more accurate (in their estimation). This has implications for model performance as well as documentation burden, as CONCERN was intentionally designed to work without adding documentation to clinicians’ workload. It also reflects a paradigm shift. Nurses are accustomed to rule-based scoring systems such as Morse Fall Risk [41] in which they enter clear assessment points to directly calculate a risk score, whereas CONCERN uses machine learning to model existing documentation without soliciting direct input from clinicians. As predictive CDSSs such as CONCERN do not involve that direct connection, nurses may require direct connections to patient outcomes or more thorough and detailed local explanations to trust predictions. Finally, only physicians used analogies to describe their understanding of CONCERN. This may be reflective of the contention by Lee [42] that humans tend to anthropomorphize goal-directed intelligent systems and may be unique to physicians in this study because CONCERN leverages nursing rather than physician documentation.

Limitations

This study has several limitations. Our use of the human-computer trust framework [17] may have biased clinicians toward certain conceptualizations. For example, we prompted clinicians to compare predictive CDSSs with expert clinicians as guided by the operationalization of perceived technical competence by Madsen and Gregor [17]. Without this prompt, clinicians may not have compared their abilities with the abilities of predictive CDSSs. Our specific questions about CONCERN (and about predictive CDSSs generally) as well as the heterogeneity in exposure to CONCERN limit our ability to know which findings are unique to CONCERN. Future research with other predictive CDSSs should control system exposure to further characterize the phenomenon of clinician trust. As with any qualitative research, our findings may not be transferable to other settings and populations. For example, perceptions may be different among older clinicians whose training and residency may not have involved EHRs and CDSSs. We also did not successfully recruit any clinicians who worked in intensive care; therefore, our findings may not be transferable to clinicians using predictive CDSSs in intensive care settings. There are also limitations inherent to remote interviews over video. We had limited ability to read nonverbal language, and 6% (1/17) of the participants opted not to turn on their camera. However, field notes were taken during the interviews, capturing tone of voice and nonverbal language. Finally, social desirability may have affected responses as the participants knew that the

interviewer was with the CONCERN team, which may have led them to self-censor negative perceptions of CONCERN.

Implications for Research, Practice, and Policy

The findings of this investigation elucidate future areas of inquiry. First, it will be important to explore the differences in requirements for trust between differing versions of predictive CDSSs. For example, CONCERN does not recommend a discrete action, whereas other systems pair predictions with a recommended action. We found that clinicians’ preferences for a recommended action varied and influenced trust. Furthermore, the extent to which clinicians rely on the predictive CDSS was shown to influence both the impact of discordant predictions and the detail needed from explanations. This may indicate that predictive CDSSs that are prescriptive or essential to the workflow will require more concordance or explanation detail than those that are informative, such as early warning systems.

It will also be important to evaluate the reception of CONCERN’s global and local explanations over sustained use and with reinforced education. It is clear from these findings that clinicians are oriented toward rule-based logic and this should be accounted for in explainable artificial intelligence research. Future research should also investigate whether nurses using CONCERN in fact change their documentation and, if so, whether those changes result in increased documentation burden or variation in predictive model performance. Finally, future work should be dedicated to investigating clinician personal attributes that may contribute to the variation in factors influencing trust.

It may be advantageous for hospital administrators to implement policies for development and implementation of predictive CDSSs aimed at increasing trust and adoption. Our findings suggest that involving clinicians in model development, allowing them to provide feedback after implementation, designing user-centered explanations for predictive CDSSs, and educating clinicians on machine learning may be effective policies for increasing trust.

Conclusions

Clinician trust in predictive CDSSs is critical for increased adoption of data-driven patient care. Our investigation of the phenomenon of clinician trust in predictive CDSSs for in-hospital deterioration produced needed knowledge on the factors that influence clinician trust. We found that perceptions of trust were largely the same between nurses and prescribing providers. Future work should investigate the relationship between perceived actionability and trust, research explanations that enhance understandability, and explore policies aimed at facilitating trust.

Acknowledgments

This work was supported by National Institute for Nursing Research grants 5T32NR007969 and 5R01NR016941-05. The authors acknowledge the clinicians who participated in their interviews and thank them for generously sharing their time and insights. The authors would also like to thank Dr Min Jeoung Kang, Graham Lowenthal, Sandy Cho, and Michelle Nghiem for their assistance in recruitment and overall support of this project. JMS was affiliated with the School of Nursing at Columbia University at the time of this work and is currently affiliated with the Department of Biomedical Informatics at Columbia University. EL

was affiliated with the Department of Biomedical Informatics at Columbia University at the time of this work and is currently affiliated with Weill Cornell Medicine.

Authors' Contributions

JMS, KDC, SCR, and PCD conceptualized the study objective. JMS collected the data. JMS, SRM, EL, and MG conducted the data analysis. MG advised on data collection and analysis. JMS wrote the initial draft of the manuscript with feedback and revisions from MG, SCR, PCD, SRM, EL, and KDC.

Conflicts of Interest

None declared.

Multimedia Appendix 1

Research materials and extended results.

[[DOCX File , 32 KB - humanfactors_v9i2e33960_appl.docx](#)]

References

1. Middleton B, Sittig DF, Wright A. Clinical decision support: a 25 year retrospective and a 25 year vision. *Yearb Med Inform* 2016 Aug 02;Suppl 1:S103-S116 [[FREE Full text](#)] [doi: [10.15265/IYS-2016-s034](#)] [Medline: [27488402](#)]
2. Tonekaboni S, Joshi S, McCradden M, Goldenberg A. What clinicians want: contextualizing explainable machine learning for clinical end use. In: *Proceedings of Machine Learning Research*. 2019 Presented at: Machine Learning for Healthcare; Aug 8-10, 2019; Ann Arbor, MI URL: <http://proceedings.mlr.press/v106/tonekaboni19a/tonekaboni19a.pdf>
3. Wang D, Yang Q, Abdul A, Lim B. Designing theory-driven user-centric explainable AI. In: *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems*. 2019 Presented at: CHI '19: CHI Conference on Human Factors in Computing Systems; May 4 - 9, 2019; Glasgow Scotland UK. [doi: [10.1145/3290605.3300831](#)]
4. Holzinger A, Biemann C, Pattichis C, Kell D. What do we need to build explainable AI systems for the medical domain? *arXiv* 2017:1-28 [[FREE Full text](#)]
5. The European Parliament and the Council of the European Union. Chapter 3 Rights of the data subject. *Official Journal of the European Union* 2016 Apr 27:39-47 [[FREE Full text](#)] [doi: [10.1093/oso/9780198826491.003.0147](#)]
6. Diprose WK, Buist N, Hua N, Thurier Q, Shand G, Robinson R. Physician understanding, explainability, and trust in a hypothetical machine learning risk calculator. *J Am Med Inform Assoc* 2020 Apr 01;27(4):592-600 [[FREE Full text](#)] [doi: [10.1093/jamia/ocz229](#)] [Medline: [32106285](#)]
7. Sandhu S, Lin AL, Brajer N, Sperling J, Ratliff W, Bedoya AD, et al. Integrating a machine learning system into clinical workflows: qualitative study. *J Med Internet Res* 2020 Nov 19;22(11):e22421 [[FREE Full text](#)] [doi: [10.2196/22421](#)] [Medline: [33211015](#)]
8. Beede E, Baylor E, Hersch F, Iurchenko A, Wilcox L, Ruamviboonsuk P, et al. A human-centered evaluation of a deep learning system deployed in clinics for the detection of diabetic retinopathy. In: *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems*. 2020 Presented at: CHI '20: CHI Conference on Human Factors in Computing Systems; Apr 25 - 30, 2020; Honolulu HI USA. [doi: [10.1145/3313831.3376718](#)]
9. Strohm L, Hehakaya C, Ranschaert ER, Boon WP, Moors EH. Implementation of artificial intelligence (AI) applications in radiology: hindering and facilitating factors. *Eur Radiol* 2020 Oct 26;30(10):5525-5532 [[FREE Full text](#)] [doi: [10.1007/s00330-020-06946-y](#)] [Medline: [32458173](#)]
10. Romero-Brufau S, Wyatt KD, Boyum P, Mickelson M, Moore M, Cognetta-Rieke C. A lesson in implementation: a pre-post study of providers' experience with artificial intelligence-based clinical decision support. *Int J Med Inform* 2020 May;137:104072. [doi: [10.1016/j.ijmedinf.2019.104072](#)] [Medline: [32200295](#)]
11. Fan W, Liu J, Zhu S, Pardalos PM. Investigating the impacting factors for the healthcare professionals to adopt artificial intelligence-based medical diagnosis support system (AIMDSS). *Ann Oper Res* 2020 Nov;294:567-592. [doi: [10.1007/s10479-018-2818-y](#)]
12. Muralitharan S, Nelson W, Di S, McGillion M, Devereaux PJ, Barr NG, et al. Machine learning-based early warning systems for clinical deterioration: systematic scoping review. *J Med Internet Res* 2021 Feb 04;23(2):e25187 [[FREE Full text](#)] [doi: [10.2196/25187](#)] [Medline: [33538696](#)]
13. Ginestra JC, Giannini HM, Schweickert WD, Meadows L, Lynch MJ, Pavan K, et al. Clinician perception of a machine learning-based early warning system designed to predict severe sepsis and septic shock. *Crit Care Med* 2019;47(11):1477-1484. [doi: [10.1097/ccm.0000000000003803](#)]
14. Giannini HM, Ginestra JC, Chivers C, Draugelis M, Hanish A, Schweickert WD, et al. A machine learning algorithm to predict severe sepsis and septic shock: development, implementation, and impact on clinical practice. *Crit Care Med* 2019 Nov;47(11):1485-1492 [[FREE Full text](#)] [doi: [10.1097/CCM.0000000000003891](#)] [Medline: [31389839](#)]

15. Jacobs M, Pradier MF, McCoy TH, Perlis RH, Doshi-Velez F, Gajos KZ. How machine-learning recommendations influence clinician treatment selections: the example of the antidepressant selection. *Transl Psychiatry* 2021 Feb 04;11(1):108 [FREE Full text] [doi: [10.1038/s41398-021-01224-x](https://doi.org/10.1038/s41398-021-01224-x)] [Medline: [33542191](https://pubmed.ncbi.nlm.nih.gov/33542191/)]
16. Zlotnik A, Alfaro MC, Pérez MC, Gallardo-Antolín A, Martínez JM. Building a decision support system for inpatient admission prediction with the Manchester triage system and administrative check-in variables. *Comput Inform Nurs* 2016 May;34(5):224-230. [doi: [10.1097/CIN.0000000000000230](https://doi.org/10.1097/CIN.0000000000000230)] [Medline: [26974710](https://pubmed.ncbi.nlm.nih.gov/26974710/)]
17. Madsen M, Gregor S. Measuring human-computer trust. In: Proceedings of the 11 th Australasian Conference on Information Systems. 2000 Presented at: Proceedings of the 11 th Australasian Conference on Information Systems; 2000; Brisbane Qld Australia URL: <http://citeseerx.ist.psu.edu/viewdoc/summary?doi=10.1.1.93.3874>
18. Sandelowski M. Whatever happened to qualitative description? *Res Nurs Health* 2000 Aug;23(4):334-340. [doi: [10.1002/1098-240x\(200008\)23:4<334::aid-nur9>3.0.co;2-g](https://doi.org/10.1002/1098-240x(200008)23:4<334::aid-nur9>3.0.co;2-g)] [Medline: [10940958](https://pubmed.ncbi.nlm.nih.gov/10940958/)]
19. Kim H, Sefcik JS, Bradway C. Characteristics of qualitative descriptive studies: a systematic review. *Res Nurs Health* 2017 Feb;40(1):23-42 [FREE Full text] [doi: [10.1002/nur.21768](https://doi.org/10.1002/nur.21768)] [Medline: [27686751](https://pubmed.ncbi.nlm.nih.gov/27686751/)]
20. Hsieh H, Shannon SE. Three approaches to qualitative content analysis. *Qual Health Res* 2005 Nov;15(9):1277-1288. [doi: [10.1177/1049732305276687](https://doi.org/10.1177/1049732305276687)] [Medline: [16204405](https://pubmed.ncbi.nlm.nih.gov/16204405/)]
21. Kang MJ, Dykes PC, Korach TZ, Zhou L, Schnock KO, Thate J, et al. Identifying nurses' concern concepts about patient deterioration using a standard nursing terminology. *Int J Med Inform* 2020 Jan;133:104016 [FREE Full text] [doi: [10.1016/j.ijmedinf.2019.104016](https://doi.org/10.1016/j.ijmedinf.2019.104016)] [Medline: [31707264](https://pubmed.ncbi.nlm.nih.gov/31707264/)]
22. Rossetti SC, Knaplund C, Albers D, Dykes PC, Kang MJ, Korach TZ, et al. Healthcare process modeling to phenotype clinician behaviors for exploiting the signal gain of clinical expertise (HPM-ExpertSignals): development and evaluation of a conceptual framework. *J Am Med Inform Assoc* 2021 Jun 12;28(6):1242-1251 [FREE Full text] [doi: [10.1093/jamia/ocab006](https://doi.org/10.1093/jamia/ocab006)] [Medline: [33624765](https://pubmed.ncbi.nlm.nih.gov/33624765/)]
23. Collins SA, Vawdrey DK. "Reading between the lines" of flow sheet data: nurses' optional documentation associated with cardiac arrest outcomes. *Appl Nurs Res* 2012 Nov;25(4):251-257 [FREE Full text] [doi: [10.1016/j.apnr.2011.06.002](https://doi.org/10.1016/j.apnr.2011.06.002)] [Medline: [22079746](https://pubmed.ncbi.nlm.nih.gov/22079746/)]
24. Collins SA, Cato K, Albers D, Scott K, Stetson PD, Bakken S, et al. Relationship between nursing documentation and patients' mortality. *Am J Crit Care* 2013 Jul;22(4):306-313 [FREE Full text] [doi: [10.4037/ajcc2013426](https://doi.org/10.4037/ajcc2013426)] [Medline: [23817819](https://pubmed.ncbi.nlm.nih.gov/23817819/)]
25. Morse JM. The significance of saturation. *Qual Health Res* 1995 May 01;5(2):147-149. [doi: [10.1177/104973239500500201](https://doi.org/10.1177/104973239500500201)]
26. MacQueen KM, McLellan E, Kay K, Milstein B. Codebook development for team-based qualitative analysis. *CAM J* 1998;10(2):31-36. [doi: [10.1177/1525822X980100020301](https://doi.org/10.1177/1525822X980100020301)]
27. Streubert HJ, Carpenter DR. *Qualitative Research in Nursing: Advancing the Humanistic Imperative*. Philadelphia, Pennsylvania, United States: Wolters Kluwer Health/Lippincott Williams & Wilkins; 2011:20-23.
28. Guba EG. Criteria for assessing the trustworthiness of naturalistic inquiries. *Educ Commun Technol* 1981 Jun;29(2):75-91. [doi: [10.1007/bf02766777](https://doi.org/10.1007/bf02766777)]
29. Spall S. Peer debriefing in qualitative research: emerging operational models. *Qual Inquiry* 1998 Jun 01;4(2):280-292. [doi: [10.1177/107780049800400208](https://doi.org/10.1177/107780049800400208)]
30. Caruana R, Lou Y, Gehrke J, Koch P, Sturm M, Elhadad N. Intelligible models for healthcare: predicting pneumonia risk and hospital 30-day readmission. In: Proceedings of the 21th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining. 2015 Presented at: KDD '15: The 21th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining; Aug 10 - 13, 2015; Sydney NSW Australia. [doi: [10.1145/2783258.2788613](https://doi.org/10.1145/2783258.2788613)]
31. Rudin C. Stop explaining black box machine learning models for high stakes decisions and use interpretable models instead. *Nat Mach Intell* 2019 May 13;1(5):206-215. [doi: [10.1038/s42256-019-0048-x](https://doi.org/10.1038/s42256-019-0048-x)]
32. Barda AJ, Horvat CM, Hochheiser H. A qualitative research framework for the design of user-centered displays of explanations for machine learning model predictions in healthcare. *BMC Med Inform Decis Mak* 2020 Oct 08;20:257 [FREE Full text] [doi: [10.1186/s12911-020-01276-x](https://doi.org/10.1186/s12911-020-01276-x)] [Medline: [33032582](https://pubmed.ncbi.nlm.nih.gov/33032582/)]
33. Sutton RA, Sharma P. Overcoming barriers to implementation of artificial intelligence in gastroenterology. *Best Pract Res Clin Gastroenterol* 2021 Jun;52-53:101732. [doi: [10.1016/j.bpg.2021.101732](https://doi.org/10.1016/j.bpg.2021.101732)] [Medline: [34172254](https://pubmed.ncbi.nlm.nih.gov/34172254/)]
34. Scheetz J, Rothschild P, McGuinness M, Hadoux X, Soyer HP, Janda M, et al. A survey of clinicians on the use of artificial intelligence in ophthalmology, dermatology, radiology and radiation oncology. *Sci Rep* 2021 Mar 04;11:5193 [FREE Full text] [doi: [10.1038/s41598-021-84698-5](https://doi.org/10.1038/s41598-021-84698-5)] [Medline: [33664367](https://pubmed.ncbi.nlm.nih.gov/33664367/)]
35. Elish MC. The stakes of uncertainty: developing and integrating machine learning in clinical care. *Ethnographic Praxis Industry Conference Proc* 2019 Jan 24;2018(1):364-380. [doi: [10.1111/1559-8918.2018.01213](https://doi.org/10.1111/1559-8918.2018.01213)]
36. Payrovnaziri SN, Chen Z, Rengifo-Moreno P, Miller T, Bian J, Chen JH, et al. Explainable artificial intelligence models using real-world electronic health record data: a systematic scoping review. *J Am Med Inform Assoc* 2020 Jul 01;27(7):1173-1185 [FREE Full text] [doi: [10.1093/jamia/ocaa053](https://doi.org/10.1093/jamia/ocaa053)] [Medline: [32417928](https://pubmed.ncbi.nlm.nih.gov/32417928/)]
37. Verghese A, Shah NH, Harrington RA. What this computer needs is a physician: humanism and artificial intelligence. *JAMA* 2018 Jan 02;319(1):19-20. [doi: [10.1001/jama.2017.19198](https://doi.org/10.1001/jama.2017.19198)] [Medline: [29261830](https://pubmed.ncbi.nlm.nih.gov/29261830/)]
38. Wiens J, Saria S, Sendak M, Ghasssemi M, Liu VX, Doshi-Velez F, et al. Do no harm: a roadmap for responsible machine learning for health care. *Nat Med* 2019 Sep;25(9):1337-1340. [doi: [10.1038/s41591-019-0548-6](https://doi.org/10.1038/s41591-019-0548-6)] [Medline: [31427808](https://pubmed.ncbi.nlm.nih.gov/31427808/)]

39. Schwartz JM, Moy AJ, Rossetti SC, Elhadad N, Cato KD. Clinician involvement in research on machine learning-based predictive clinical decision support for the hospital setting: a scoping review. *J Am Med Inform Assoc* 2021 Mar 01;28(3):653-663 [FREE Full text] [doi: [10.1093/jamia/ocaa296](https://doi.org/10.1093/jamia/ocaa296)] [Medline: [33325504](https://pubmed.ncbi.nlm.nih.gov/33325504/)]
40. Cabitza F, Rasoini R, Gensini GF. Unintended consequences of machine learning in medicine. *JAMA* 2017 Aug 08;318(6):517-518. [doi: [10.1001/jama.2017.7797](https://doi.org/10.1001/jama.2017.7797)] [Medline: [28727867](https://pubmed.ncbi.nlm.nih.gov/28727867/)]
41. Morse JM, Morse RM, Tylko SJ. Development of a scale to identify the fall-prone patient. *Can J Aging* 1989;8(4):366-377. [doi: [10.1017/s0714980800008576](https://doi.org/10.1017/s0714980800008576)]
42. Lee SS. Philosophical evaluation of the conceptualisation of trust in the NHS' Code of Conduct for artificial intelligence-driven technology. *J Med Ethics* 2022 Apr 16;48(4):272-277. [doi: [10.1136/medethics-2020-106905](https://doi.org/10.1136/medethics-2020-106905)] [Medline: [33593871](https://pubmed.ncbi.nlm.nih.gov/33593871/)]

Abbreviations

CDSS: clinical decision support system

CONCERN: Communicating Narrative Concerns Entered by Registered Nurses

EHR: electronic health record

PA: physician assistant

RN: registered nurse

Edited by A Kushniruk; submitted 30.09.21; peer-reviewed by Z Yin, J Xiao, J Bagby; comments to author 20.01.22; revised version received 02.03.22; accepted 21.03.22; published 12.05.22.

Please cite as:

Schwartz JM, George M, Rossetti SC, Dykes PC, Minshall SR, Lucas E, Cato KD

Factors Influencing Clinician Trust in Predictive Clinical Decision Support Systems for In-Hospital Deterioration: Qualitative Descriptive Study

JMIR Hum Factors 2022;9(2):e33960

URL: <https://humanfactors.jmir.org/2022/2/e33960>

doi: [10.2196/33960](https://doi.org/10.2196/33960)

PMID: [35550304](https://pubmed.ncbi.nlm.nih.gov/35550304/)

©Jessica M Schwartz, Maureen George, Sarah Collins Rossetti, Patricia C Dykes, Simon R Minshall, Eugene Lucas, Kenrick D Cato. Originally published in *JMIR Human Factors* (<https://humanfactors.jmir.org>), 12.05.2022. This is an open-access article distributed under the terms of the Creative Commons Attribution License (<https://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in *JMIR Human Factors*, is properly cited. The complete bibliographic information, a link to the original publication on <https://humanfactors.jmir.org>, as well as this copyright and license information must be included.

Original Paper

A Web-Based Stratified Stepped Care Mental Health Platform (TourHeart): Semistructured Interviews With Stakeholders

Emily W S Tsoi¹, PhD; Winnie W S Mak¹, PhD; Connie Y Y Ho¹, MSSc; Gladys T Y Yeung², MSSc

¹Department of Psychology, The Chinese University of Hong Kong, Shatin, NT, China (Hong Kong)

²New Life Psychiatric Rehabilitation Association, Kowloon, China (Hong Kong)

Corresponding Author:

Winnie W S Mak, PhD

Department of Psychology, The Chinese University of Hong Kong

Room 354 Sino Building

Shatin, NT

China (Hong Kong)

Phone: 852 39436577

Fax: 852 26035019

Email: wsmak@cuhk.edu.hk

Abstract

Background: TourHeart, a web-based stratified stepped care mental health platform, is a one-stop solution that integrates psychoeducation and other well-being promotional tools for mental health promotion and mental illness prevention and evidence-based, low-intensity psychological interventions for the treatment of people with anxiety and depressive symptoms. Instead of focusing only on symptom reduction, the platform aims to be person-centered and recovery-oriented, and continual feedback from stakeholders is sought. Understanding the perspectives of users and service providers enables platform developers to fine-tune both the design and content of the services for enhanced service personalization and personal recovery.

Objective: This qualitative study evaluated a web-based mental health platform by incorporating the perspectives of both users and service providers who administered the platform and provided coaching services. The platform included both web-based and offline services targeting adults along the mental health spectrum based on the two-continua model of mental health and mental illness.

Methods: Interview questions were designed based on the Reach, Effectiveness, Adoption, Implementation, and Maintenance framework (RE-AIM). Views on offline services, the design of the web-based platform, user experience, and the contents of the platform were explored using semistructured interviews. A total of 27 service users and 22 service providers were recruited using purposive criterion sampling. A hybrid thematic analysis was performed to identify salient aspects of users' and providers' experiences with and views of the platform.

Results: Totally, 3 broad themes (namely, the quality of the platform, drivers for platform use, and coaching services) emerged from the interview data that highlighted users' views of and experiences with the web-based platform. The platform's general esthetics, operations, and contents were found to be critical features and drivers for continued use. Although coaching services were indispensable, participants preferred the autonomy and anonymity associated with web-based mental health services.

Conclusions: This study highlights the importance of web-based mental health services being easy to navigate and understand, being user-centric, and providing adequate guidance in self-help. It also confirms existing design standards and recommendations and suggests that more rigorous, iterative user experience research and robust evaluation should be conducted in the future adaptation of web-based stratified stepped care services, so that they can be more personalized and better promote personal recovery.

(*JMIR Hum Factors* 2022;9(2):e35057) doi:[10.2196/35057](https://doi.org/10.2196/35057)

KEYWORDS

stepped care; recovery; engagement; eHealth; mental health; mental well-being; psychological intervention

Introduction

Background

This study aimed to evaluate a web-based stratified stepped care mental health platform for the working adult population. Under the stratified stepped care approach, the platform recommends services that are commensurate with the intensity needed by users based on the results of a brief self-report mental health assessment on anxiety, depression, and flourishing that was completed at the beginning and then provides timely step-down or step-up options to users upon completion of interventions at the recommended level and based on their latest mental health status. This approach aims to optimize the provision of services based on users' needs and efficient use of resources, given that the mental health of the working adult population is often overlooked. Specifically, mental ill-health is a tremendous burden on the society owing to loss of productivity and quality of life among the workforce. A meta-analysis [1] on the global prevalence of common mental disorders in the general adult population (aged 16–65 years) indicated a global common mental disorders prevalence of 17.5% (in the past 12 months). The statistics vary across geographical locations. Among all high-income countries, individuals from English-speaking countries (eg, United States and United Kingdom) reported the highest prevalence (19%), followed by European countries (17.1%) and Asian countries (11.5%). A report from the World Health Organization in 2017 revealed that the total number of people living with depression is 322 million worldwide, and nearly half of these people live in the South-East Asia region and Western Pacific region [2]. Despite the substantial social and economic impact of mental ill-health and its associated loss of health and functioning on individuals and the society, help seeking, especially in Asian countries, remains low. Commonly cited barriers to help seeking include stigma, long waiting time, high costs, and concerns about privacy and anonymity [3,4].

In addition to mental ill-health, whether the general population is flourishing is also an important concern, as it impacts the extent to which individuals can attain complete mental health [5]. Under the two-continua model of mental health and mental illness, any person can fall within the orthogonal quadrants, from languishing to flourishing in the dimension of mental health and from no mental illness to having mental illness in the dimension of mental illness [6]. Thus, in addition to paying attention to the prevention and treatment of mental illness, promotion of mental health is equally essential at the population level to maximize public mental health and personal recovery [7].

Given the growing concerns about the accessibility, affordability, scalability, and anonymity of mental health promotion and help seeking for mental illness, digital mental health apps and platforms have burgeoned as robust solutions in recent years. A review of the literature [8] identified several digital mental health service types, including psychoeducation, screening and assessment, social support, guided self-help, and intervention.

For digital interventions, evidence has been established regarding their effectiveness in promoting well-being, reducing

mental illness and stigma around help seeking, and ameliorating stress and psychological distress [9–11]. Typically, digital interventions are designed based on a single treatment approach such as mindfulness-based training [12,13], cognitive behavioral therapy (CBT) [14,15], or a combination of multiple treatment approaches [16], although increasing amount of research has suggested transdiagnostic approaches to be effective in tackling underlying mechanisms for common mental disorders such as anxiety and depression [17].

Although stratified care can provide timely services to individuals that are commensurate with their mental health needs, stepped care service begins with the least restrictive care to individuals to maximize resources so that users with varying levels of needs are provided the same intensity of services and may progress to services of high intensity on an as-needed basis. Its accessibility and flexibility in stepping up and down are particularly suitable for promoting mental health and reducing mental illness within the same web-based platform for a population with varying levels of flourishing, anxiety, and depression. Through mobile technology, different levels of services can be tailored and delivered to individuals who would otherwise not pay attention to their mental health or receive timely services, as their conditions are usually considered as less urgent and less acute than others with more severe mental health needs. People with low levels of mental health needs are also found to be more hesitant in seeking help for their issues [18].

TourHeart—Web-Based Stratified Stepped Care Platform

This study aimed to evaluate a stratified stepped care service platform with both offline and web-based components (The Jockey Club TourHeart Project) that was designed based on the matched care and stepped care principles [19,20]. The web-based platform targets adults on a spectrum, from flourishing with normal to severe range of anxiety and depressive symptoms. Through a simple sign-in procedure, users are directed to a web-based personal assessment in which their mental health profiles are obtained. At the time of writing, 9738 individuals have registered, signed in, and completed the battery of mental health profiles at least once. On the basis of their initial scores for flourishing (assessed by the Flourishing Scale) [21], anxiety symptoms (assessed by the 7-item Generalized Anxiety Disorder questionnaire) [22], and depressive symptoms (assessed by the 9-item Patient Health Questionnaire) [23,24], users are recommended services that are commensurate with their current mental health state based on 4 levels (see Table 1 for the service-level criteria). Level 1 targets the general public and includes social media and public events such as psychoeducational talks and public exhibitions that promote mental health awareness and reduce mental illness stigma (ie, offline events). Level 2 provides self-guided web-based mental well-being promotion training programs for improving well-being and reducing psychological distress for those who present with mild levels of anxiety or depressive symptoms across the flourishing spectrum. Web-based courses for this level include stress management, emotion regulation training, self-compassion training, and mindfulness-based training. Each training course consists of 4 modules, and

participants are invited to complete 1 module each week sequentially. Level 3 provides coach-guided web-based psychological interventions for alleviation of anxiety and depressive symptoms based on guided self-help principles for those who present with moderate levels of anxiety and depressive symptoms across the flourishing spectrum. At this level, users can access mindfulness-based intervention or rumination-focused CBT (RFCBT) [25]. These interventions contain 6 sequential modules, with each module requiring 1 week of commitment to complete. Users are guided at the end of each module by feedback on their written homework and exercises, through email from well-being coaches (psychology graduates with specific training in low-intensity psychological interventions). Finally, level-4 services are low-intensity psychological interventions for those who present with

moderately severe to severe levels of anxiety and depressive symptoms across the flourishing spectrum. Participants at this level receive individual sessions of low-intensity CBT (LiCBT) delivered by trained psychological well-being officers in person or via videoconferencing. Figure 1 shows the stratified stepped care model for mental well-being used in TourHeart. Owing to the complexity and fluidity of the platform where multiple components may interact with users' experiences and thus affect users' outcomes [26], qualitative method is particularly useful in encapsulating the perspectives of multiple stakeholders. Therefore, service users and service providers (event organizers, project administrators, and well-being coaches) at all levels were invited to participate in this study to provide their views on the platform.

Table 1. Service levels and criteria^a.

Level	PHQ-9 ^b score, range	Combination logic	GAD-7 ^c score, range	Combination logic	Suicide risk (PHQ-9—score for question 9=2 to 3)
1	N/A ^d	N/A	N/A	N/A	N/A
2 (nonclinical to mild)	0-9	AND	0-7	AND	No
3 (moderate)	10-14	OR	8-14	AND	No
4 (moderately severe to severe)	15-27	OR	15-21	OR	Yes
5 (complex mental health issues)	— ^e	—	—	—	—

^aFlourishing is assessed for all platform users. Levels 2 to 4 are reserved for individuals who have registered on the platform, whereas level 1 is targeted to all working populations.

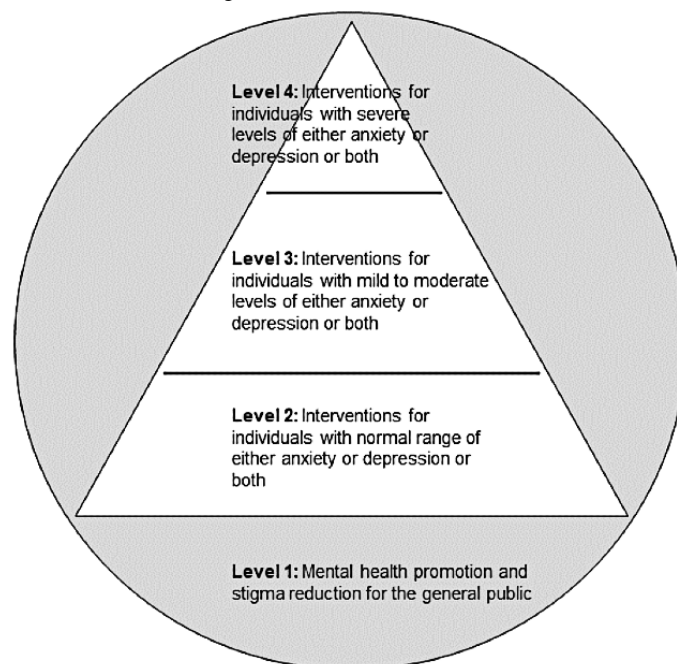
^bPHQ-9: 9-item Patient Health Questionnaire.

^cGAD-7: 7-item Generalized Anxiety Disorder questionnaire.

^dN/A: not applicable.

^eNot available (manually assigned and referred out).

Figure 1. Stratified stepped care model for mental well-being.



Aims and Objectives

Given the importance of understanding the contents and salient features of the platform that may exert positive effects on its users to improve future design and user experience, this study aimed to explore the views and preferences of users and service providers regarding the implementation of the TourHeart platform with both offline and web-based services targeted for working adults along the spectra of flourishing, anxiety, and depressive symptoms. The specific study objectives were to (1) explore users' views on the design and quality of web-based and offline mental health services and (2) identify drivers for web-based platform use.

Methods

This qualitative study used semistructured interviews that were conducted via phone or videoconferencing with both users and service providers to explore the experiences and perceived usefulness of the services and the platform. The technical and design aspects of the web-based platform were also explored.

Ethics Approval

This study obtained ethics approval from the corresponding author's (WWSM) home institution (Clinical Research Ethics Committee 2018.654) and followed the Standards for Reporting Qualitative Research [27].

Recruitment

To maximize the variation in experiences obtained from our participants, purposive criterion sampling [28] was used. A list of potential participants was generated from the back end of the platform. Participants who have accessed TourHeart for any of the 4 levels or have taken courses from levels 2 to 4 in the past 4 weeks from the time of the study were identified as potential participants. The target sample size was 6 participants from level 1, 8 participants each from levels 2 and 3, and 6 participants from level 4. Half of the participants from levels 2 to 4 (10/21, 48%) were *completers* and the remaining participants were *dropouts*, defined as those who attempted >30% but <100% of the course.

Participants were recruited on a rolling basis until the target sample size was achieved. For the user group, potential participants were invited via WhatsApp messages and a follow-up phone call to explain the purpose and implications of joining the research in detail. A total of 58 participants were invited, of whom 21 (36%) declined or did not respond. Then, the consenting participants (27/58, 47%) were invited to complete a consent form via the web, where their basic demographic information was also collected. After obtaining consent, phone interviews were scheduled. Each interview lasted for an average of 40 minutes.

The service provider group (n=22) consisted of both past and current research, administrative, and service staff. They were invited either in person or via WhatsApp messages. Consenting staff were invited to complete a web-based consent form, through which basic demographic information was also collected. To increase synergy between study participants and to obtain detailed responses, separate focus groups were

conducted based on the roles of the participants. The grouping of roles included (1) project administrators (project manager and project administrative assistant), (2) past research staff, (3) current research staff and well-being coaches, and (4) volunteers who organize offline events. Each group consisted of 2 staff members.

All participants, except the current project staff, received a HK \$100 (US \$12.75) honorarium in the form of cash as a token of appreciation for the time spent for the study.

Interviews

The interview protocol was designed based on the RE-AIM (Reach, Effectiveness, Adoption, Implementation, and Maintenance) framework [29], focusing on areas that were most relevant to the participants' needs and experiences with the web-based platform. In the users' version, aspects related to reach, effectiveness, and implementation were explored using open-ended questions. In the service providers' version, aspects related to reach, effectiveness, adoption, service implementation, and future maintenance were explored. Equipped with previous knowledge and experiences with the web-based platform, the interviewer used probes where appropriate; digressions were allowed for obtaining organic responses. All conversations were recorded and transcribed verbatim.

Data Analysis

An initial codebook was developed by the 2 interviewers using a hybrid approach (inductive and deductive), followed by thematic analytic procedure [30]. Thematic analysis typically consists of 6 phases. They include (1) familiarizing with the raw data, (2) initial coding, (3) searching for themes, (4) reviewing the themes, (5) defining and naming themes, and (6) reporting of findings. These 6 phases do not have to be followed in a sequential manner. To ensure data integrity, triangulating questions were asked during all the interviews to confirm the researcher's own understanding of the participants' views; in addition, notes were taken during each interview. Moreover, immediately after each interview, brief analytic summaries were made to document the researcher's personal observations, such as tone of the participant's voice. These steps were taken to ensure that reporting remained as faithful to the original narratives as possible. NVivo (version 12; QSR International) was used to conduct the analysis.

Results

Participants

Most participants in the user group (n=27) were women (22/27, 82%), with a mean age of 34 (SD 10.11; range 19-58) years. Most of these participants (16/27, 59%) reported to have obtained at least tertiary-level education. Participants came from diverse work backgrounds. One-third of the participants were employed in the education sector (9/27, 33%), followed by those employed in public health and welfare sector (4/27, 15%), banking (3/27, 11%), and public services (3/27, 11%). The remaining participants were from other industries. Regarding the positions held, of the 70% (19/27) of the participants who opted to provide information about their work, 47% (9/19) participants reported to be holding executive positions and an

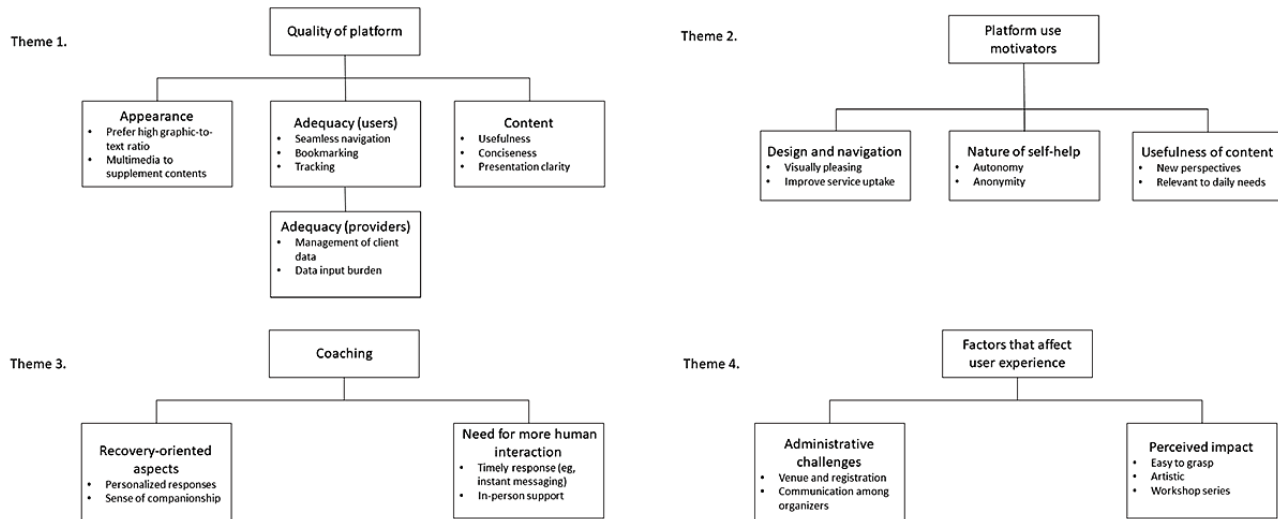
equal number of the remaining participants reported to be office workers (5/19, 26%) or homemakers and carers (5/19, 26%) at the time of the interview. Finally, most of those from the service provider group were women (18/22, 82%), with a mean age of 34 (SD 10.27; range 21-46) years.

Overview of Key Findings

This study focused on evaluating both the offline and web-based services offered by the TourHeart platform. For evaluation of offline services, two subthemes emerged from one broad theme (*factors that influence user experience*), which were (1)

administrative challenges and (2) perceived impact. For the web-based services, findings pertinent to the positive and desired features of the platform that can inform future development and implementation were analyzed and reported. With this analytical goal in mind, codes from 21 user interviews and 14 service provider interviews were grouped into three major themes: (1) quality of the platform, (2) motivating factors for platform use, and (3) comments on human coaching. These themes were divided into subthemes. The themes and an overview of associated findings are summarized in Figure 2.

Figure 2. Themes and subthemes.



Web-Based Services

Theme 1: Quality of the Platform

Overview

This part of the study involved 60% (21/35) users (level 2: 8/21, 38%; level 3: 8/21, 38%; and level 4: 5/21, 24%) and 40% (14/35) providers. At the beginning of the interview, participants were invited to freely share their experience of using the platform to identify critical platform design parameters perceived by its users. Responses from this segment of interviews were coded into 3 main subthemes, namely, appearance, adequacy, and content of the platform.

Appearance

In this subtheme, comments on the graphics, graphics-text balance, use of multimedia, style consistency, word length, language and style, and color choices were collated.

The platform adopted a minimalist design with pastel tone colors (Figures 3 and 4). A total of 38% (8/21) of the users commented

the website design to be attractive and “very eye pleasing” in terms of its choice of colors and the overall layout. Users generally preferred high graphic-to-text ratio. Specifically, the use of multimedia to supplement course content was mentioned frequently (6/21, 29% users). The type of media included animation, videos, voice narration, infographics, and minigames that could facilitate their understanding of concepts introduced in the courses, such as the topic on *thinking traps* that was introduced in the RFCBT course. Users cited the following:

People these days are generally very busy and those who visit the platform are not in the best of mood. Therefore, the platform could use a bit more videos or animations to illustrate their points. No one wants to read a whole page filled with words. [Connie, level 3]

Voice narration is good, or at least more illustrations...I don't like to read as I am already constantly tired.... [Ivy, level 3]

Figure 3. Screenshot of the landing page for the TourHeart platform.

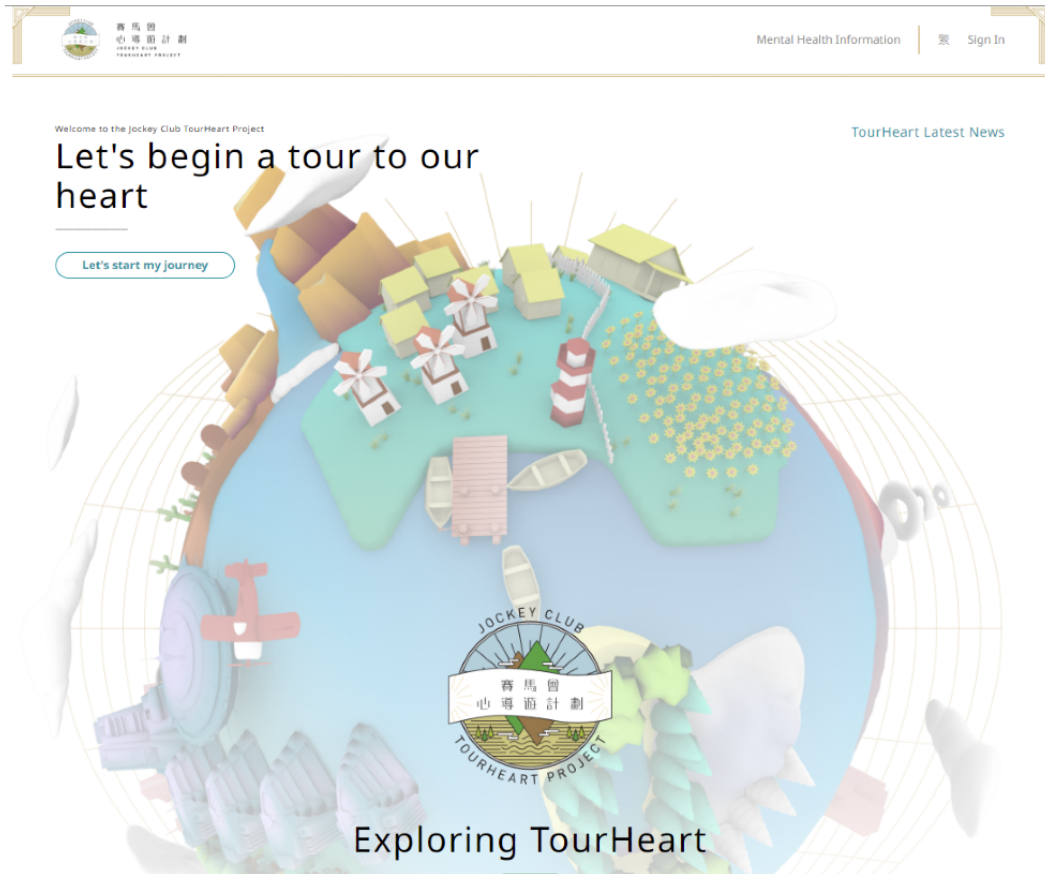
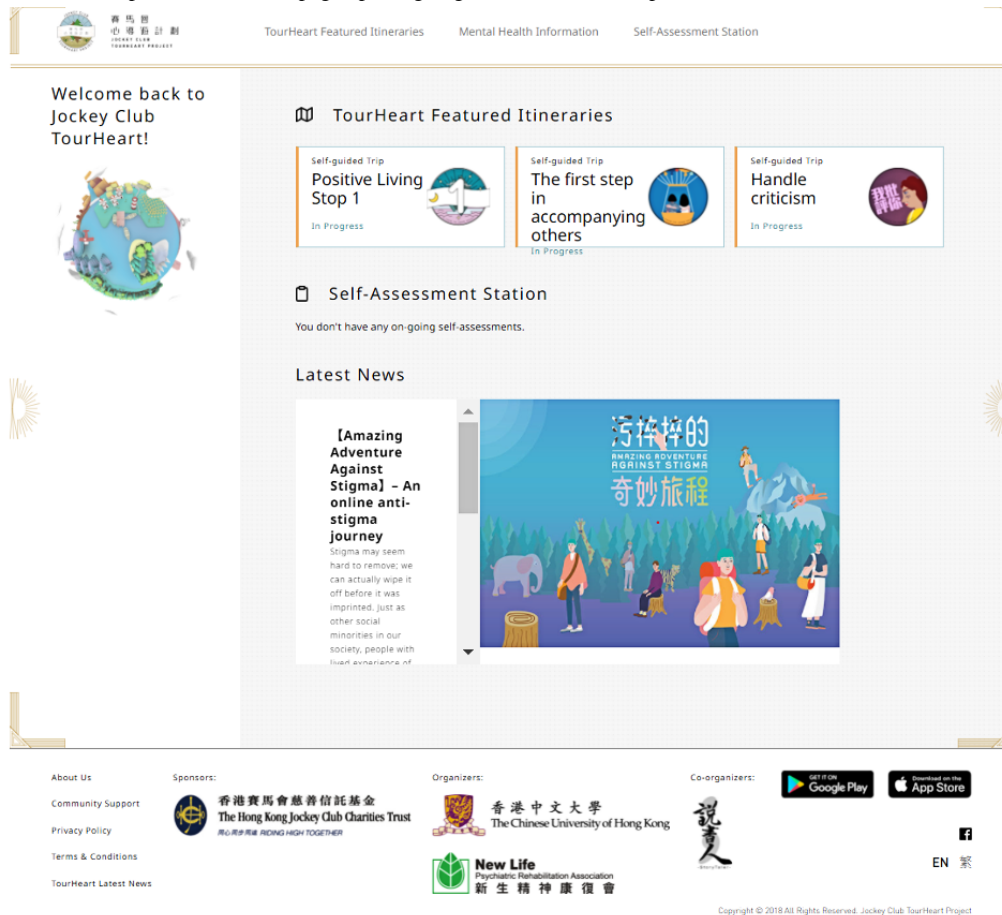


Figure 4. Screenshot of a user's personalized front page upon signing in to the TourHeart platform.



Adequacy

All service users were asked about the technical adequacy of the platform (21/21, 100%); most service users had either positive or no comments about the technical adequacy of the platform (14/21, 67%). The remaining 33% (7/21) of the users commented on various technical inadequacies of the platform. In general, critical comments pertained to ease of navigation, reliability, bookmarking function, and interactivity. Of these 7 users, 5 (71%) users had difficulty with the bookmarking function. A loss of bookmark was experienced after logging out from their sessions. For instance, courses did not resume from points where they were stopped and the users had to restart at different points. Similarly, users reported the audio being switched off when the phone's screen was in sleep mode.

Moreover, some users were not entirely accustomed to the self-guided nature of the platform. On the TourHeart platform, users were free to choose courses that were made accessible to them at their level of mental health states, such as mindfulness-based training and RFCBT, but some users thought the provision of options was confusing:

I felt confused at the beginning. When I started using the platform, two courses were made available to me, one was mindfulness-related and the other was CBT-related. I remember thinking...which course should I take? Both were available but I felt confused as to where and how I should start. [Kennis, level 3]

In terms of suggestions, 10% (2/21) of the users expressed that they would be interested in functions where courses could have different versions of varying lengths as an option for users to choose from for more flexibility, and another 10% (2/21) of the users stated that they would like to have an additional feature where course contents could be archived for offline access. Finally, 5% (1/21) of the users suggested the use of pop-up messages after signing in to draw attention to timely topics.

During the interviews with service providers and project staff (14/35, 40%), comments were primarily about the limitations of the platform back end to properly manage user data, such as limitations in searching and filtering. These comments were grouped into three categories: (1) issues related to user data management (14/14, 100%), (2) data input burden as a deterrent to adherence (4/14, 29%), and (3) limited human resources available owing to the high level of demand for coaching (2/14, 14%). Illustrative quotes pertaining to these 3 categories, respectively, are listed in the following sections.

User Management

Service providers and project staff reported a few platform functions that are not user-friendly or missing functions that may expedite and streamline their work in user management.

We had difficulty using the search function to filter out users that fit under specific categories. Quite often we had to use their email that they used for registration; but a lot of users have more than one emails and sometimes we cannot locate a user within the system. The system can be improved with a more sophisticated filtering function.

We had to type in progress notes manually, with no preset categories. For example, age of onset, current medications. This makes progress tracking somewhat difficult. So, we had to set up our own templates using another application.

As part of the intervention, we had to make progress notes for each user after each session. The problem with this function is that it cannot be edited after I clicked submitted. Therefore, if I made a typo or if I had amendment that I want to make to the old progress note, I had no choice but to submit a second one.

The single biggest difficulty was that we spent a lot of our working hours sending out email reminders to users when we could put our time into other uses. We had to send emails manually to guide users every step of the way. For example, welcoming emails to those who newly registered; reminders to those who have not visited the platform for a certain period of time; or emails of encouragement to those who have completed a certain milestone. Indeed, we had to manually create excel sheets to keep track of movements of all the users in the platform.

Since we need to send out so many emails manually, we are left with little time. So, we had to use templates when we gave feedback to the users. If we had more time, we could have done more and done better.

Data Input Burden

Service providers and project staff also noted that some users may be turned off by the need to respond to assessment questions prior to starting a course, which may explain user attrition.

From the backend, we could see that users terminated themselves without beginning of a course. We guess they might be scared off by the long assessment questions (24 questions) that they were required to answer before being assigned to the course(s).

Content

Another prominent subtheme that arose from the interview discussions was about the courses offered and whether they were useful for the participants in managing their psychological distress. Only participants who received services from levels 2 and 3 (16/21, 76%) commented on this topic. As discussion of the specific components of the courses and whether they were useful in alleviating users' distress is beyond the scope of this paper, we collated comments that illustrated the general quality of the contents, namely (1) usefulness of contents (16/16, 100%), (2) conciseness (6/16, 38%), and (3) presentation clarity (4/16, 25%).

The contents offered on the platform were generally well received for their effectiveness in facilitating users to deal with emotional vicissitudes for themselves and others. For instance, users said the following:

I have learned to take care of myself emotionally, and through this, take care of my better half, and my

family. This is the inspiration I have. [Edward, level 3]

I used to have ruminative thoughts, especially after a bad day at work. Now, instead of ruminating about the negative experiences at work, I learned to be mindful when I walk home. This helped me to feel more settled. [Ivy, level 3]

Regarding conciseness of the content, users seemed to prefer to have fewer words or even organize the contents into shorter sections or even point forms:

I think they could break down the contents into point forms, as some people might go through the website as if it is a tool book. [Joyce, level 2]

I wanted to capture the feedbacks from coach; however, they seemed to be too long and it takes time for me to revisit all of the materials. I would prefer to read the materials in point-form.... [Phillip, level 3]

Regarding presentation clarity, participants commented that the contents were at appropriate level of difficulty and presented in a way that can be easily understood:

I first learned of mindfulness from the platform. I also read books about mindfulness and found that the information and audio tracks contained are adequate for my daily usage. [Yan, level 2]

Theme 2: Platform Use Motivators

Overview

This part of the study involved 59% (16/27) of the users (level 2: 8/16, 50% and level 3: 8/16, 50%). An open-ended question was asked to solicit participants' motivations for using the platform. The responses are divided into a few categories in the order of frequency; they include (1) design and navigation (6/16, 38%), (2) the nature of self-help (5/16, 31%), (3) usefulness of contents (4/16, 25%), and (4) free of charge (1/16, 6%).

Design and Navigation

Despite encountering challenges in using the platform, the convenience of the web-based platform with its pleasant design was the most frequently mentioned motivating factor that attracted users:

The webpage looks young and vibrant. Secondly, the use of graphics makes it less dull to look at. The design is great; I remember the animated earth on the sign-in page, it gave me a strong impression. I like it. [Rosemary, level 3]

It is very convenient, there is no need for any pen and paper. [Joyce, level 2]

Nature of Self-help

The autonomy and anonymity associated with web-based self-help were considered to be the second most significant motivating factor in platform use. Some of the comments include the following:

The best thing...everything is under my control. It is self-help and the self-assessment enabled me to learn about my current state. It is good. [Priscilla, level 2]

I can log-in any time I like, without having to arrange any appointment. [Ivy, level 3]

I like the concept of self-help; it is amongst the first of its kind that is available in Hong Kong. [Sunny, level 3]

I really enjoyed using it in my own private time. It is less embarrassing compared to in-person interventions. I try to avoid meeting real people for therapies. [Cecilia, level 2]

Usefulness of Content

Perceived usefulness and satisfaction with the contents offered were the third most frequently mentioned factors that affected platform use. The platform was said to offer new perspectives and information that are helpful to users' mental well-being:

I find the well-being tips to be useful. I remember one that asked us to smile even though we were not in a good mood, contrary to my expectations, the tip actually worked for me. [Irene, level 2]

I feel calm just by looking at the materials. [Cecilia, level 2]

The contents offered here are very rich, with audios and teachings about mindfulness...body scan has been particularly helpful for me when I am having sleepless nights. [Joyce, level 2]

The part about worries was particularly useful for me. Given the current situation in Hong Kong, local news made me feel distressed and helpless. So, I applied what I learned from the course – I approach or challenge my worries systematically, by asking myself: are my worries constructive? Is there anything I can do to alleviate or distract myself from it? I also learn to communicate my worries with others as now I understand that worrying is normal. [Sunny, level 3]

Theme 3: Human Coaching

Overview

Totally, 30% (8/27; level 3) of users provided feedback about the coaching service that they received as part of their intervention at level 3. All the responses from participants were positive about the recovery-oriented messages they received. Some of them expressed the need for more human interactions in support of using the web-based self-help platform.

Recovery-Oriented Aspects

All the users were satisfied with the quality of coaching received through emails. The empowering, person-centered, and recovery-oriented language used by the coaches in the email communications were viewed as "thoughtful," which also gave them "a sense of companionship." Totally, 25% (2/8) of the users said that the coaching component was the primary motivator for platform use:

I was impressed with the long and detailed replies from the coaches. I can see that they have poured in a lot of effort to give me feedback and address my concerns. The interactivity of coaching I received from this platform is different from other websites where I mostly accessed just to get information. The collaborative nature of this platform is very useful for me. [Rosemary, level 3]

I know that they have read all of the things I have written and provided responses tailored to my needs and worries...the coaches are like my companion. [Kennis, level 3]

Need for More Human Interactions

When we asked these 8 users for suggestions to improve the coaching service, 5 (63%) of them preferred to have more interactions with the coaches in a timely manner. Of this 63% (5/8) of the users, 60% (3/5) users preferred to have instant messaging with a human coach, 20% (1/5) preferred to have regular phone conversations, and the remaining 20% (1/5) preferred regular face-to-face meetings with coaches.

Offline Services

Overview

This part of the study involved 43% (6/14; level 1) users and 57% (8/14; event organizers) service providers. Data from the semistructured interviews were analyzed to identify topics that were mentioned frequently by the participants. The topics stated by both users and providers who received and delivered level-1 (offline) services, respectively, were collated under 1 general theme: *factors that influence users' experience* of activities that were part of the promotion and prevention services. Topics under each theme are reported according to their frequencies.

Theme 4: Factors That Influence Users' Experience of Offline Service

Administrative Challenges

This part of the study involved 50% (6/12; level 1) users and 50% (6/12) volunteers. Our data indicated that venues should be easily accessible to the public and properly equipped for comfort (6/12, 50%). As most people in Hong Kong do not own a car, the location of the venue must be convenient and close to a subway station (Mass Transit Railway). In addition, air-conditioned venues should be chosen for events held in the summer, as participants from the public exhibition indicated that they were not able to stay and explore the entire exhibition (held indoor without air-conditioning) owing to the summer heat.

Furthermore, 25% (3/12) of the participants mentioned that an effective registration system was crucial. Features such as immediate notification of registration outcome, update of waitlist status, and reminder before the event using instant messaging were mentioned. These features were said to be particularly important for people with busy schedules.

Among event organizers, a few (4/6, 67%) mentioned that the biggest administrative challenge was working together as a team because they all come from diverse professional backgrounds.

All of these volunteers (4/4, 100%) suggested regular team-building exercises to foster better work relationships.

Perceived Impact

This part of the study involved 65% (11/17) users (level 1: 6/11, 55% and level 4: 5/11, 45%) and 35% (6/17) volunteers. Most users and providers (10/17, 59%) gave positive comments about the public mental health promotion events centered on the theme *accompany*. These workshops highlighted the core concept of how everyone can be both the supporter and recipient of support, which reduces stigma of help seeking. During these events, participants were also equipped with skills to identify the more obvious signs of depression and anxiety to promote awareness within themselves and among others. In general, participants found all the events to be very "down-to-earth" (Jasmine, level 1) and the concepts to be easy for laypeople to grasp. The intended messages of the psychoeducational workshop and exhibition were also said to be delivered effectively and thoughtfully (10/17, 59%). For instance, the event organizers, some of whom worked closely with people with lived experience to deliver talks or *storytelling sessions*, found that sharing of lived experience within small groups was particularly impactful in building connections and strengthening empathy between people with lived experience and the audience. In addition, in the public exhibition, participants were invited to inflate balloons to reveal messages printed on them. During the process, participants had to exert patience and effort, similar to the qualities required to be good listeners, which was the key message of that exhibition. The psychoeducation talks were also said to "leave a lasting impression" (Ivy, level 1) that "normalizes mental illnesses" (Harry, level 1) without overwhelming lay participants. However, at the same time, some participants also expressed a wish to receive more in-depth knowledge or skills-based training. For example, some suggested that the organizer conduct a series of workshops progressing through the details of various mental health-related content so that participants can acquire the full range of skills on various aspects of mental health management.

Regarding level 4, users commented on LiCBT (face-to-face or via videoconferencing). Of the 5 level-4 users, 1 (20%) user decided to drop out of the service as she preferred talk therapy such as counseling. The rest of the users (4/5, 80%) reported positive experience with the service, including feeling a sense of warmth and rapport with the psychological well-being officer (5/5, 100%) and finding the guidebooks used in LiCBT as useful (3/5, 60%). Notably, the presence and interactions between the psychological well-being officer and user were mentioned by all users as the most positive aspect of level-4 service, marking the human element indispensable in both our web-based and offline services:

It's the sense of being understood...techniques aside...when I talk to the PWO, I felt that she was present and was willing to listen...that already made me feel so much better. [Eugene, level 4]

Discussion

Principal Findings

Overall, our findings highlighted the importance of person-centered design that emphasizes autonomy and competence, with relevant information empowering users to self-help. The implementation of a web-based stratified stepped care approach to mental health support is driven by the need for increased accessibility and improved availability of relevant mental health services to the general working population. Although studies in the field have demonstrated the effectiveness of digital mental health interventions [9-11], their uptake is still moderately suboptimal [31]. Moreover, an integrated web-based platform curating relevant mental health tools and interventions for promotion, prevention, and treatment of common mental health conditions is lacking. With the recent development of web-based psychotherapies that use artificial intelligence and chatbots, novel technology-based therapeutic experiences with better user experience are being increasingly recognized.

This qualitative study was designed to evaluate a stratified stepped care platform that includes both offline and web-based mental health services. Primarily, the desired and positive qualities of the TourHeart platform were explored from the perspectives of service users and providers. Factors that were identified to be important in improving users' experience in offline services include administrative challenges and perceived impact. In the evaluation of web-based components of the platform, 3 broad themes emerged from the interview data: quality of the platform, drivers for platform use, and human coaching. Across the themes, a visually attractive platform with high graphic-to-text ratio and multimedia content, seamless user experience in navigation with a sophisticated bookmarking function that can cater to users who may be using the services in between their schedules and thus might log off intermittently, and concise down-to-earth presentation of relevant content emerged as the most important. These qualities are consistent with previous research that shows platform's esthetics to be facilitative of initial uptake [32,33]. Thus, these are critical design parameters that developers should focus on while developing future web-based platforms. In addition to esthetics and ease of use, users also preferred the sense of autonomy and anonymity that is associated with the web-based platform [34,35] along with alleviation of distress that drives their continued use. In particular, many people may regard using a web-based platform as less embarrassing and more preferred than conventional face-to-face therapies.

Although human interactions via email coaching (level 3) and via face-to-face or videoconferencing with psychological well-being officers (level 4) are considered to be indispensable components of the platform, some contradictions were noted. Specifically, although self-help and the resulting sense of autonomy and mastery were found to improve overall use, several participants found self-directed navigation of low-intensity psychological interventions to be challenging, as they had to pace and tailor the materials by themselves without support from other people. Another contradictory finding was the need for human interactions to support platform use. These

contradictory findings are broadly consistent with the findings from previous meta-synthesis [36] that proposed 2 critical system characteristics to address the competing needs of users. The first characteristic was collaboration between users and the system. This sense of collaboration should be different from having a sense of *enforced autonomy*, where users are burdened with having to navigate the system without any type of external support. The second characteristic was connection that can be achieved by balancing the need for personal privacy with increased distance and maximizing interactions with therapists or web-based coaches through messaging, forum discussions, emails, and so on.

In summary, findings of our study are consistent with existing literature and design standards for digital mental health services [37,38] and highlight the importance of services to be both user-centric and comprehensible to all users. In addition to these front-end features, a back-end system that is effective for system administrators to manage user data is also essential for effective service delivery, including functions for searching, filtering, and reporting. Notably, themes generated from our study were found to converge with the objective domains covered in the Mobile Application Rating Scale [39], namely, users' engagement, functionality, esthetics, and information quality, which supports the validity of our findings.

Regarding future directions, findings of this study suggested that developers should integrate more motivational and interactive components to improve users' engagement, such as statements of encouragement and support, open-ended and fill-in-the-blank entries for users to reflect on the related questions and apply course exercises to themselves in their daily lives. In addition to improving initial uptake, the general esthetics of the web-based platform can also facilitate users' adherence [32,33]. Moreover, clear and concise information is needed to provide a seamless self-help experience for users. Our data indicated that a lack of understanding about the stratified stepped care service model might render the user experience moderately unclear. A lack of road map in guiding the course of treatment and manual data input were 2 factors that hampered users' overall use. Therefore, service expectations of users should be taken into account in the design phase to achieve better alignment. Owing to the dynamic nature of stepped care service where users are subjected to stepping up and down to interventions that vary in intensity and interactivity in accordance with their self-reported mental health status, changes to the types or length of interventions could be confusing to some users. In fact, evidence from another qualitative study conducted by the authors of this paper (EWS Tsoi, PhD, unpublished data, 2021) indicated that some users are ambivalent to this approach to care. Specifically, several users with high levels of distress were unwilling to be stepped up and receive face-to-face care. These participants preferred to continue with web-based care to maintain their anonymity. The fear of disclosure and feelings of embarrassment during face-to-face therapies might deter some individuals from stepping up to face-to-face interventions. Future web-based services should consider tailoring to the needs of this particular group of users using blended web-based–offline approach or

web-based psychological interventions that are suitable for people with more severe levels of anxiety or depression.

Although technical struggles were not the focus of our study, the difficulties encountered by our participants and their preferences for support in platform use suggested that support function that is instant and highly accessible should be included. In recent years, text-based conversational agents (eg, chatbot) have emerged to provide timely and reliable customer service experience in fields such as banking and web-based shopping [40,41]. Chatbots have also been increasingly adopted in health and mental health care settings [42-44]. Although its use and effectiveness are still largely experimental, how this new technology could be leveraged is a productive area to explore. Ideally, the involvement of real persons to provide support behind the screen should be kept to a minimum so that services can be scalable and delivered more cost-effectively to more users [45]. Future studies can further explore ways in which paraprofessionals and peer support workers can deliver timely support remotely.

TourHeart is operated under the stratified stepped care model that is similar to Stepped Care 2.0, which focuses more on user experience and recovery in a progressive manner [46]. Furthermore, TourHeart covers the spectrum of mental health needs, which includes mental health promotion, stigma reduction, mental illness prevention, and treatment for common mental disorders. The second phase of TourHeart (Jockey Club TourHeart+ Project) models on Stepped Care 2.0 by using a person-centered approach and machine learning, where users can choose options at different levels and receive regular feedback and reports on growth or deterioration so that they can retain a sense of autonomy. Iterative user experience research was also conducted during the course of platform development and maintenance. By understanding the perspectives of various stakeholders, web-based mental health platforms may be better designed in the future to provide timely and bespoke web-based

and offline mental health-related information, skills, and interventions to users, with technical support and consideration of user experience to improve the acceptability and effectiveness of digital mental health services.

Strengths and Limitations

Using a bottom-up analytical approach, this study has identified salient aspects of the web-based self-help platform, TourHeart, that were important to both users and service providers and were the main drivers of platform use. In addition, our findings highlighted the importance of balancing our need for support and need for privacy and autonomy, which could potentially be addressed through enhanced system collaboration and connections, which were explained in the previous section. Factors that promote engagement can also be further explored through user experience research.

This study was limited by the heterogeneity of the sample in terms of age, gender, and clinical representation. Participants who agreed to participate in the study may also pose a risk of self-selection bias. In addition, comments on the platform in general were invited and participants without in-depth knowledge of user experience and design may not be able to conceptualize what an *ideal* platform should be; hence, the desired features of inadequacies pointed out should be interpreted with some caution.

Conclusions

This study identified a few issues that need to be addressed to enhance the adoption and usability of web-based mental health platforms. These insights were incorporated into the development of phase 2 of TourHeart (TourHeart+). Given that people's needs and preferences for mental health services may evolve over time, incorporation of iterative and rigorous user experience research into the development and maintenance of web-based mental health self-help platforms is urgently needed.

Acknowledgments

The Jockey Club TourHeart Project was funded by the Hong Kong Jockey Club Charities Trust (reference number: 2017/0070).

Conflicts of Interest

None declared.

References

1. Steel Z, Marnane C, Iranpour C, Chey T, Jackson JW, Patel V, et al. The global prevalence of common mental disorders: a systematic review and meta-analysis 1980-2013. *Int J Epidemiol* 2014 Apr;43(2):476-493 [FREE Full text] [doi: [10.1093/ije/dyu038](https://doi.org/10.1093/ije/dyu038)] [Medline: [24648481](https://pubmed.ncbi.nlm.nih.gov/24648481/)]
2. Depression and Other Common Mental Disorders: Global Health Estimates. World Health Organization. URL: <https://apps.who.int/iris/bitstream/handle/10665/254610/WHO-MSD-MER-2017.2-eng.pdf> [accessed 2022-04-11]
3. Lam LC, Wong CS, Wang MJ, Chan WC, Chen EY, Ng RM, et al. Prevalence, psychosocial correlates and service utilization of depressive and anxiety disorders in Hong Kong: the Hong Kong Mental Morbidity Survey (HKMMS). *Soc Psychiatry Psychiatr Epidemiol* 2015 Sep;50(9):1379-1388. [doi: [10.1007/s00127-015-1014-5](https://doi.org/10.1007/s00127-015-1014-5)] [Medline: [25660760](https://pubmed.ncbi.nlm.nih.gov/25660760/)]
4. Tong AC, Tsoi EW, Mak WW. Socioeconomic status, mental health, and workplace determinants among working adults in Hong Kong: a latent class analysis. *Int J Environ Res Public Health* 2021 Jul 26;18(15):7894 [FREE Full text] [doi: [10.3390/ijerph18157894](https://doi.org/10.3390/ijerph18157894)] [Medline: [34360188](https://pubmed.ncbi.nlm.nih.gov/34360188/)]

5. Keyes CL. Promoting and protecting mental health as flourishing: a complementary strategy for improving national mental health. *Am Psychol* 2007;62(2):95-108. [doi: [10.1037/0003-066X.62.2.95](https://doi.org/10.1037/0003-066X.62.2.95)] [Medline: [17324035](https://pubmed.ncbi.nlm.nih.gov/17324035/)]
6. Keyes CL. The mental health continuum: from languishing to flourishing in life. *J Health Soc Behav* 2002 Jun;43(2):207-222. [Medline: [12096700](https://pubmed.ncbi.nlm.nih.gov/12096700/)]
7. Provencher HL, Keyes CL. Complete mental health recovery: bridging mental illness with positive mental health. *J Public Ment Health* 2011 Mar;10(1):57-69 [FREE Full text] [doi: [10.1108/17465721111134556](https://doi.org/10.1108/17465721111134556)]
8. Lal S, Adair CE. E-mental health: a rapid review of the literature. *Psychiatr Serv* 2014 Jan 01;65(1):24-32. [doi: [10.1176/appi.ps.201300009](https://doi.org/10.1176/appi.ps.201300009)] [Medline: [24081188](https://pubmed.ncbi.nlm.nih.gov/24081188/)]
9. Spijkerman MP, Pots WT, Bohlmeijer ET. Effectiveness of online mindfulness-based interventions in improving mental health: a review and meta-analysis of randomised controlled trials. *Clin Psychol Rev* 2016 Apr;45:102-114 [FREE Full text] [doi: [10.1016/j.cpr.2016.03.009](https://doi.org/10.1016/j.cpr.2016.03.009)] [Medline: [27111302](https://pubmed.ncbi.nlm.nih.gov/27111302/)]
10. Stratton E, Lampit A, Choi I, Calvo RA, Harvey SB, Glozier N. Effectiveness of eHealth interventions for reducing mental health conditions in employees: a systematic review and meta-analysis. *PLoS One* 2017 Dec 21;12(12):e0189904 [FREE Full text] [doi: [10.1371/journal.pone.0189904](https://doi.org/10.1371/journal.pone.0189904)] [Medline: [29267334](https://pubmed.ncbi.nlm.nih.gov/29267334/)]
11. Phillips E, Gordeev VS, Schreyögg J. Effectiveness of occupational e-mental health interventions: a systematic review and meta-analysis of randomized controlled trials. *Scand J Work Environ Health* 2019 Nov 01;45(6):560-576 [FREE Full text] [doi: [10.5271/sjweh.3839](https://doi.org/10.5271/sjweh.3839)] [Medline: [31184758](https://pubmed.ncbi.nlm.nih.gov/31184758/)]
12. Cillessen L, van de Ven MO, Compen FR, Bisseling EM, van der Lee ML, Speckens AE. Predictors and effects of usage of an online mindfulness intervention for distressed cancer patients: usability study. *J Med Internet Res* 2020 Oct 02;22(10):e17526 [FREE Full text] [doi: [10.2196/17526](https://doi.org/10.2196/17526)] [Medline: [33006567](https://pubmed.ncbi.nlm.nih.gov/33006567/)]
13. Sun Y, Li Y, Wang J, Chen Q, Bazzano AN, Cao F. Effectiveness of smartphone-based mindfulness training on maternal perinatal depression: randomized controlled trial. *J Med Internet Res* 2021 Jan 27;23(1):e23410 [FREE Full text] [doi: [10.2196/23410](https://doi.org/10.2196/23410)] [Medline: [33502326](https://pubmed.ncbi.nlm.nih.gov/33502326/)]
14. Titov N, Dear BF, Staples LG, Bennett-Levy J, Klein B, Rapee RM, et al. MindSpot clinic: an accessible, efficient, and effective online treatment service for anxiety and depression. *Psychiatr Serv* 2015 Oct;66(10):1043-1050. [doi: [10.1176/appi.ps.201400477](https://doi.org/10.1176/appi.ps.201400477)] [Medline: [26130001](https://pubmed.ncbi.nlm.nih.gov/26130001/)]
15. Fitzpatrick K, Darcy A, Vierhile M. Delivering cognitive behavior therapy to young adults with symptoms of depression and anxiety using a fully automated conversational agent (Woebot): a randomized controlled trial. *JMIR Ment Health* 2017 Jun 06;4(2):e19 [FREE Full text] [doi: [10.2196/mental.7785](https://doi.org/10.2196/mental.7785)] [Medline: [28588005](https://pubmed.ncbi.nlm.nih.gov/28588005/)]
16. Cillessen L, Schellekens MP, Van de Ven MO, Donders AR, Compen FR, Bisseling EM, et al. Consolidation and prediction of long-term treatment effect of group and online mindfulness-based cognitive therapy for distressed cancer patients. *Acta Oncol* 2018 Oct;57(10):1293-1302. [doi: [10.1080/0284186X.2018.1479071](https://doi.org/10.1080/0284186X.2018.1479071)] [Medline: [29932784](https://pubmed.ncbi.nlm.nih.gov/29932784/)]
17. Newby JM, McKinnon A, Kuyken W, Gilbody S, Dalgleish T. Systematic review and meta-analysis of transdiagnostic psychological treatments for anxiety and depressive disorders in adulthood. *Clin Psychol Rev* 2015 Aug;40:91-110 [FREE Full text] [doi: [10.1016/j.cpr.2015.06.002](https://doi.org/10.1016/j.cpr.2015.06.002)] [Medline: [26094079](https://pubmed.ncbi.nlm.nih.gov/26094079/)]
18. Cross SP, Hickie I. Transdiagnostic stepped care in mental health. *Public Health Res Pract* 2017 Apr 27;27(2):2721712 [FREE Full text] [doi: [10.17061/phrp2721712](https://doi.org/10.17061/phrp2721712)] [Medline: [28474049](https://pubmed.ncbi.nlm.nih.gov/28474049/)]
19. O'Donohue WT, Draper C. The case for evidence-based stepped care as part of a reformed delivery system. In: Draper C, O'Donohue WT, editors. *Stepped Care and e-Health*. New York, NY, USA: Springer; Oct 7, 2010:1-16.
20. van Straten A, Tiemens B, Hakkaart L, Nolen WA, Donker MC. Stepped care vs. matched care for mood and anxiety disorders: a randomized trial in routine practice. *Acta Psychiatr Scand* 2006 Jun;113(6):468-476. [doi: [10.1111/j.1600-0447.2005.00731.x](https://doi.org/10.1111/j.1600-0447.2005.00731.x)] [Medline: [16677223](https://pubmed.ncbi.nlm.nih.gov/16677223/)]
21. Diener E, Wirtz D, Tov W, Kim-Prieto C, Choi D, Oishi S, et al. New well-being measures: short scales to assess flourishing and positive and negative feelings. *Soc Indic Res* 2009 May 28;97(2):143-156. [doi: [10.1007/s11205-009-9493-y](https://doi.org/10.1007/s11205-009-9493-y)]
22. Spitzer RL, Kroenke K, Williams JB, Löwe B. A brief measure for assessing generalized anxiety disorder: the GAD-7. *Arch Intern Med* 2006 May 22;166(10):1092-1097. [doi: [10.1001/archinte.166.10.1092](https://doi.org/10.1001/archinte.166.10.1092)] [Medline: [16717171](https://pubmed.ncbi.nlm.nih.gov/16717171/)]
23. Kroenke K, Spitzer RL, Williams JB. The PHQ-9: validity of a brief depression severity measure. *J Gen Intern Med* 2001 Sep;16(9):606-613 [FREE Full text] [doi: [10.1046/j.1525-1497.2001.016009606.x](https://doi.org/10.1046/j.1525-1497.2001.016009606.x)] [Medline: [11556941](https://pubmed.ncbi.nlm.nih.gov/11556941/)]
24. Wang W, Bian Q, Zhao Y, Li X, Wang W, Du J, et al. Reliability and validity of the Chinese version of the Patient Health Questionnaire (PHQ-9) in the general population. *Gen Hosp Psychiatry* 2014;36(5):539-544. [doi: [10.1016/j.genhosppsy.2014.05.021](https://doi.org/10.1016/j.genhosppsy.2014.05.021)] [Medline: [25023953](https://pubmed.ncbi.nlm.nih.gov/25023953/)]
25. Watkins ER. *Rumination-Focused Cognitive-Behavioral Therapy for Depression*. New York, NY, USA: The Guilford Press; 2018.
26. Craig P, Dieppe P, Macintyre S, Michie S, Nazareth I, Petticrew M, Medical Research Council Guidance. Developing and evaluating complex interventions: the new Medical Research Council guidance. *BMJ* 2008 Sep 29;337:a1655 [FREE Full text] [doi: [10.1136/bmj.a1655](https://doi.org/10.1136/bmj.a1655)] [Medline: [18824488](https://pubmed.ncbi.nlm.nih.gov/18824488/)]
27. O'Brien BC, Harris IB, Beckman TJ, Reed DA, Cook DA. Standards for reporting qualitative research: a synthesis of recommendations. *Acad Med* 2014 Sep;89(9):1245-1251 [FREE Full text] [doi: [10.1097/ACM.0000000000000388](https://doi.org/10.1097/ACM.0000000000000388)] [Medline: [24979285](https://pubmed.ncbi.nlm.nih.gov/24979285/)]

28. Patton MQ. *Qualitative Research & Evaluation Methods: Integrating Theory and Practice*. 4th edition. Thousand Oaks, CA, USA: SAGE Publications; Nov 2014.
29. Glasgow RE, Vogt TM, Boles SM. Evaluating the public health impact of health promotion interventions: the RE-AIM framework. *Am J Public Health* 1999 Sep;89(9):1322-1327. [doi: [10.2105/ajph.89.9.1322](https://doi.org/10.2105/ajph.89.9.1322)] [Medline: [10474547](https://pubmed.ncbi.nlm.nih.gov/10474547/)]
30. Braun V, Clarke V. Using thematic analysis in psychology. *Qual Res Psychol* 2006 Jan;3(2):77-101 [FREE Full text] [doi: [10.1191/1478088706qp063oa](https://doi.org/10.1191/1478088706qp063oa)]
31. Waller R, Gilbody S. Barriers to the uptake of computerized cognitive behavioural therapy: a systematic review of the quantitative and qualitative evidence. *Psychol Med* 2009 May;39(5):705-712. [doi: [10.1017/S0033291708004224](https://doi.org/10.1017/S0033291708004224)] [Medline: [18812006](https://pubmed.ncbi.nlm.nih.gov/18812006/)]
32. Kuosmanen T, Fleming TM, Barry MM. Using computerized mental health programs in alternative education: understanding the requirements of students and staff. *Health Commun* 2018 Jun;33(6):753-761. [doi: [10.1080/10410236.2017.1309620](https://doi.org/10.1080/10410236.2017.1309620)] [Medline: [28426253](https://pubmed.ncbi.nlm.nih.gov/28426253/)]
33. Olf M. Mobile mental health: a challenging research agenda. *Eur J Psychotraumatol* 2015 May 19;6:27882 [FREE Full text] [doi: [10.3402/ejpt.v6.27882](https://doi.org/10.3402/ejpt.v6.27882)] [Medline: [25994025](https://pubmed.ncbi.nlm.nih.gov/25994025/)]
34. Anderson K, Burford O, Emmerton L. Mobile health apps to facilitate self-care: a qualitative study of user experiences. *PLoS One* 2016 May 23;11(5):e0156164 [FREE Full text] [doi: [10.1371/journal.pone.0156164](https://doi.org/10.1371/journal.pone.0156164)] [Medline: [27214203](https://pubmed.ncbi.nlm.nih.gov/27214203/)]
35. Tataru N, Arsand E, Skrøvseth SO, Hartvigsen G. Long-term engagement with a mobile self-management system for people with type 2 diabetes. *JMIR Mhealth Uhealth* 2013 Mar 27;1(1):e2432 [FREE Full text] [doi: [10.2196/mhealth.2432](https://doi.org/10.2196/mhealth.2432)] [Medline: [25100649](https://pubmed.ncbi.nlm.nih.gov/25100649/)]
36. Knowles SE, Toms G, Sanders C, Bee P, Lovell K, Rennick-Egglestone S, et al. Qualitative meta-synthesis of user experience of computerised therapy for depression and anxiety. *PLoS One* 2014 Jan 17;9(1):e84323 [FREE Full text] [doi: [10.1371/journal.pone.0084323](https://doi.org/10.1371/journal.pone.0084323)] [Medline: [24465404](https://pubmed.ncbi.nlm.nih.gov/24465404/)]
37. Bernard R, Sabariego C, Cieza A. Barriers and facilitation measures related to people with mental disorders when using the web: a systematic review. *J Med Internet Res* 2016 Jun 09;18(6):e157 [FREE Full text] [doi: [10.2196/jmir.5442](https://doi.org/10.2196/jmir.5442)] [Medline: [27282115](https://pubmed.ncbi.nlm.nih.gov/27282115/)]
38. Rotondi AJ, Eack SM, Hanusa BH, Spring MB, Haas GL. Critical design elements of e-health applications for users with severe mental illness: singular focus, simple architecture, prominent contents, explicit navigation, and inclusive hyperlinks. *Schizophr Bull* 2015 Mar;41(2):440-448 [FREE Full text] [doi: [10.1093/schbul/sbt194](https://doi.org/10.1093/schbul/sbt194)] [Medline: [24375458](https://pubmed.ncbi.nlm.nih.gov/24375458/)]
39. Stoyanov SR, Hides L, Kavanagh DJ, Wilson H. Development and validation of the user version of the mobile application rating scale (uMARS). *JMIR Mhealth Uhealth* 2016 Jun 10;4(2):e72 [FREE Full text] [doi: [10.2196/mhealth.5849](https://doi.org/10.2196/mhealth.5849)] [Medline: [27287964](https://pubmed.ncbi.nlm.nih.gov/27287964/)]
40. Quah JT, Chua YW. Chatbot assisted marketing in financial service industry. In: *Proceedings of the 16th International Conference on Services Computing*. 2019 Presented at: SCC '19; June 25-30, 2019; San Diego, CA, USA p. 107-114 URL: https://doi.org/10.1007/978-3-030-23554-3_8 [doi: [10.1007/978-3-030-23554-3_8](https://doi.org/10.1007/978-3-030-23554-3_8)]
41. Solis-Quispe JM, Quico-Cauti KM, Ugarte W. Chatbot to simplify customer interaction in e-commerce channels of retail companies. In: *Proceedings of the 2021 International Conference on Information Technology & Systems*. 2021 Jan 31 Presented at: ICITS '21; February 4-6, 2021; Libertad , Ecuador p. 561-570 URL: https://doi.org/10.1007/978-3-030-68285-9_52 [doi: [10.1007/978-3-030-68285-9_52](https://doi.org/10.1007/978-3-030-68285-9_52)]
42. Denecke K, Vaahesan S, Arulnathan A. A mental health chatbot for regulating emotions (SERMO) - concept and usability test. *IEEE Trans Emerg Topics Comput* 2021 Jul 1;9(3):1170-1182 [FREE Full text] [doi: [10.1109/tetc.2020.2974478](https://doi.org/10.1109/tetc.2020.2974478)]
43. Grové C. Co-developing a mental health and wellbeing chatbot with and for young people. *Front Psychiatry* 2021 Feb 1;11:606041 [FREE Full text] [doi: [10.3389/fpsy.2020.606041](https://doi.org/10.3389/fpsy.2020.606041)] [Medline: [33597898](https://pubmed.ncbi.nlm.nih.gov/33597898/)]
44. Tudor Car L, Dhinakaran DA, Kyaw BM, Kowatsch T, Joty S, Theng YL, et al. Conversational agents in health care: scoping review and conceptual analysis. *J Med Internet Res* 2020 Aug 07;22(8):e17158 [FREE Full text] [doi: [10.2196/17158](https://doi.org/10.2196/17158)] [Medline: [32763886](https://pubmed.ncbi.nlm.nih.gov/32763886/)]
45. Feijt MA, de Kort YA, Bongers IM, IJsselsteijn WA. Perceived drivers and barriers to the adoption of eMental health by psychologists: the construction of the levels of adoption of eMental health model. *J Med Internet Res* 2018 Apr 24;20(4):e153 [FREE Full text] [doi: [10.2196/jmir.9485](https://doi.org/10.2196/jmir.9485)] [Medline: [29691215](https://pubmed.ncbi.nlm.nih.gov/29691215/)]
46. Cornish P. *Stepped Care 2.0: A Paradigm Shift in Mental Health*. Cham, Switzerland: Springer; 2020.

Abbreviations

CBT: cognitive behavioral therapy

LiCBT: low-intensity cognitive behavioral therapy

RE-AIM: Reach, Effectiveness, Adoption, Implementation, and Maintenance

RFCBT: rumination-focused cognitive behavioral therapy

Edited by A Kushniruk; submitted 18.11.21; peer-reviewed by H van Marwijk, G Strudwick; comments to author 20.01.22; revised version received 08.03.22; accepted 21.03.22; published 13.05.22.

Please cite as:

Tsoi EWS, Mak WWS, Ho CYY, Yeung GTY

A Web-Based Stratified Stepped Care Mental Health Platform (TourHeart): Semistructured Interviews With Stakeholders

JMIR Hum Factors 2022;9(2):e35057

URL: <https://humanfactors.jmir.org/2022/2/e35057>

doi: [10.2196/35057](https://doi.org/10.2196/35057)

PMID: [35560109](https://pubmed.ncbi.nlm.nih.gov/35560109/)

©Emily W S Tsoi, Winnie W S Mak, Connie Y Y Ho, Gladys T Y Yeung. Originally published in JMIR Human Factors (<https://humanfactors.jmir.org>), 13.05.2022. This is an open-access article distributed under the terms of the Creative Commons Attribution License (<https://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Human Factors, is properly cited. The complete bibliographic information, a link to the original publication on <https://humanfactors.jmir.org>, as well as this copyright and license information must be included.

Original Paper

#BingeDrinking—Using Social Media to Understand College Binge Drinking: Qualitative Study

Madison N Cirillo¹, MPH; Jennifer P Halbert¹, MSPH; Jessica Gomez Smith², BA; Nour Sami Alamiri¹, MPH; Karen S Ingersoll², PhD

¹Department of Public Health Sciences, University of Virginia School of Medicine, Charlottesville, VA, United States

²Department of Psychiatry and Neurobehavioral Sciences, University of Virginia School of Medicine, Charlottesville, VA, United States

Corresponding Author:

Karen S Ingersoll, PhD

Department of Psychiatry and Neurobehavioral Sciences

University of Virginia School of Medicine

560 Ray C Hunt Dr

Charlottesville, VA, 22903

United States

Phone: 1 434 243 0581

Email: KES7A@hscmail.mcc.virginia.edu

Abstract

Background: Hazardous drinking among college students persists, despite ongoing university alcohol education and alcohol intervention programs. College students often post comments or pictures of drinking episodes on social media platforms.

Objective: This study aimed to understand one university's student attitudes toward alcohol use by examining student posts about drinking on social media platforms and to identify opportunities to reduce alcohol-related harm and inform novel alcohol interventions.

Methods: We analyzed social media posts from 7 social media platforms using qualitative inductive coding based on grounded theory to identify the contexts of student drinking and the attitudes and behaviors of students and peers during drinking episodes. We reviewed publicly available social media posts that referenced alcohol, collaborating with undergraduate students to select their most used platforms and develop locally relevant search terms; all posts in our data set were generated by students associated with a specific university. From the codes, we derived themes about student culture regarding alcohol use.

Results: In total, 1151 social media posts were included in this study. These included 809 Twitter tweets, 113 Instagram posts, 100 Greekrank posts, 64 Reddit posts, 34 College Confidential posts, 23 Facebook posts, and 8 YouTube posts. Posts included both implicit and explicit portrayals of alcohol use. Across all types of posts reviewed, positive drinking attitudes were most common, followed by negative and then neutral attitudes, but valence varied by platform. Posts that portrayed drinking positively received positive peer feedback and indicate that drinking is viewed by students as an essential and positive part of university student culture.

Conclusions: Social media provide a real-time picture of students' behavior during their own and others' heavy drinking. Posts portray heavy drinking as a normal part of student culture, reinforced by peers' positive feedback on posts. Interventions for college drinking should help students manage alcohol intake in real time, provide safety information during alcohol use episodes, and raise student awareness of web-based privacy concerns and reputation management. Additional interventions for students, alumni, and parents are needed to address positive attitudes about and traditions of drinking.

(*JMIR Hum Factors* 2022;9(2):e36239) doi:[10.2196/36239](https://doi.org/10.2196/36239)

KEYWORDS

college students; binge drinking; social media; young adults

Introduction

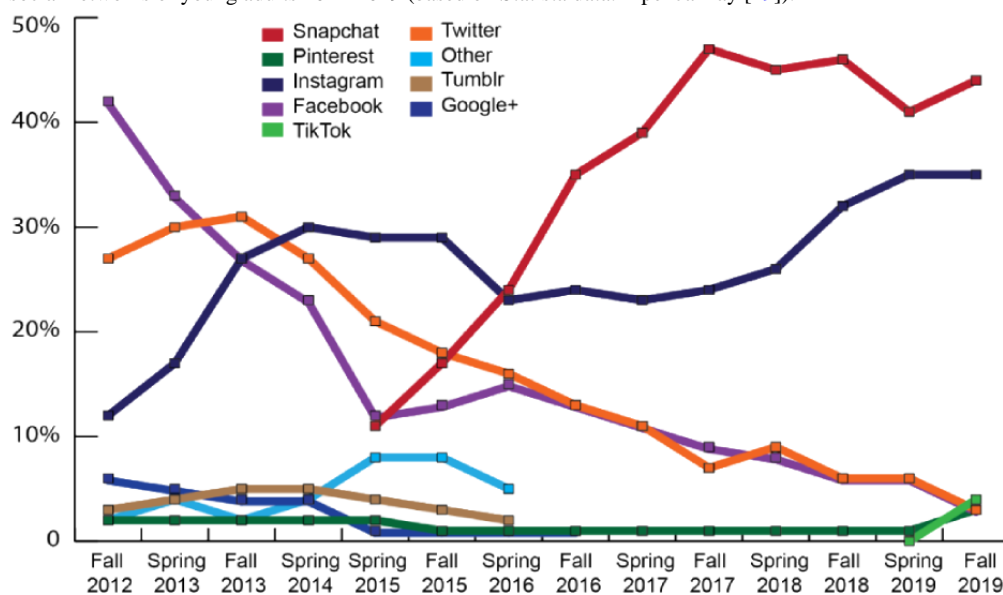
Background

Binge drinking among college students remains common and consequential. Approximately one-third of college students binge drink [1]. A quarter of college students report missing class, falling behind in class, doing poorly in examinations, and receiving lower grades due to drinking [2,3]. College students in the United States who drank in the last year reported episodes of forgetting where they were (28%), doing something they regretted (23%), blacking out (15%), having unprotected sex (14%), and injuring themselves (8%) [4]. Short-term severe effects of binge drinking include fatalities, motor vehicle crashes, poor academic achievement, and risky behaviors. Alcohol-related deaths per 100,000 students increased by 18% between 1998 and 2014, primarily because deaths due to alcohol poisoning doubled [5]. Nationwide, 11.2% of full-time college students reported symptoms indicating alcohol use disorder [6]. Longer-term effects of college binge drinking include unhealthy alcohol use after college, particularly among fraternity members whose frequent binge drinking continued through the age of 35 years [7]. College binge drinking remains a significant public

health concern despite near-universal university interventions. Novel interventions are needed to prevent tragic outcomes during the college years and persistent problems beyond them [8-10].

Social media can provide real-time behavioral data for large populations [10,11]. For example, Twitter data have been used to track influenza symptoms, estimate alcohol sales, measure depression, track HIV prevalence, and track heart disease mortality [12-18]. Various social media platforms allow for different types of expression, and platform popularity waxes and wanes over time. However, despite the favored platform changing over time (Figure 1, based on Statista research [19]), nearly 90% of people aged 18 to 29 years use social media, with most social media users posting on multiple platforms. Among Instagram users, 90% use Facebook and 50% use Twitter concurrently [20]. Although student drinking is typically measured through surveys, social media are potentially a richer source of contextual data about drinking [21]. Autobiographical social media posts contain timely thoughts, emotions, and behaviors. Users express their identities and communicate with their peers. College students portray their current alcohol use in posts about their everyday lives, interactions with others, and engagement in activities on social media platforms.

Figure 1. Favorite social networks of young adults 2012-2019 (based on Statista data: Piper Jaffray [19]).



Social media posts by college students often show risky drinking behaviors and the college drinking culture. Posts establish a web-based identity that normalizes and glamorizes binge drinking [22,23]. Most college students are exposed to peers' alcohol-related behaviors in real life or on social media. Posting about substance use on social media is nearly universal among college students [24-26]. A study of 71 profiles found that 99% had alcohol-related content, and 10% of students posted about illicit substance use [26]. Another study analyzed *drunk* tweets and their geolocation, correlated with self-reported alcohol consumption; college towns had more tweets about excessive drinking than cities and rural areas [27].

Peer influences predict drinking among college-aged populations [28]. Viewing alcohol-related social media content is significantly associated with increased risky drinking cognition,

alcohol use, and negative consequences [25,29-34]. Students usually receive positive reinforcement on the web for posting alcohol-related content [35-38]. Social validation of alcohol posts increases the number and intensity of students' drinking behavior displays over time [10,39]. College students' substance use posts are associated with self-reported outcomes [26,35,40,41]. Students with a web-based *drinking persona* have more motivation and intentions to drink [42], engage in hazardous drinking, and experience more alcohol-related consequences than peers without web-based drinking personas [43,44].

Objective

Previous studies of college students' alcohol posts have assessed single social media sites where users build social capital through

networks, such as Facebook, Instagram, or Twitter. There are additional college-focused social media sites that host anonymously shared content, such as Greekrank and College Confidential. Reddit hosts *subreddits*, topic-specific discussions, including those on specific universities. This study is the first to use data from multiple social media sites to characterize alcohol-related posts among college students. Our expanded scope is important because students typically maintain multiple social media profiles [20]. The purpose of this qualitative study was to analyze social media posts about college binge drinking from multiple platforms at one university and to identify the real-time social contexts of student drinking. We hope to provide new information about behavioral targets for intervention that could inform the development of tailored mobile health interventions for college binge drinking.

Methods

Study Design

To develop a new intervention to prevent student binge drinking at a mid-Atlantic public university, we assessed student drinking behavior and attitudes to capture aspects of drinking culture at this particular university. This university is frequently ranked in the top 50 *party schools* in the nation [45-48], and over the last decade, it has experienced alcohol-related incidents that have created national headlines. Among several substudies to inform development of this intervention, in this study, we collected, organized, coded, and analyzed social media posts using qualitative research methods. These apply a systematic process of coding content and identifying themes and patterns into classes. The results describe the context of participants' social media personas in a subjective but systematic scientific manner [49].

Data Sources

Research assistants (RAs) were undergraduate and postgraduate students who captured social media posts. Social media platforms, search terms, hashtags, and specific locations for data collection were defined through iterative discussions with current undergraduates about events, traditions, and local establishments related to the drinking culture at the university. The RAs used hashtags, locations, and search terms commonly associated with the university. RAs searched social media posts from Facebook, Twitter, Instagram, YouTube, College Confidential, Greekrank, and Reddit. These platforms were selected based on the popularity of the platform among students, popularity during the timeframe under review, ability to see posts without a connection to a poster (eg, friend or follower), and durability of content (eg, content is not time limited as on a platform such as Snapchat). Each platform has varying capabilities and levels of anonymity that determine the nature of the messages produced, as described in [Multimedia Appendix 1](#).

Data Collection and Data Exclusion

Between March 2019 and November 2019, RAs gathered publicly available posts and comments about college students' drinking behaviors published on 7 social media platforms between March 2013 and February 2019; all posts in our data

set were generated by students associated with a specific university. RAs were assigned a site, and spreadsheets were developed with the date and time, username, gender (if known), course of study (if known), search terms, likes, shares, views, comments, and content keywords for each post. Given that we used a manual search method and manually reviewed each search term, our goal was to obtain a representative sample, acknowledging that we would not be able to find all available public posts. Instagram and Facebook required log-ins, so RAs created new accounts from which they could capture these posts. To not violate the implicit trust granted by establishing a relationship with a social media poster, these profiles were for log-in purposes only; RAs did not publish any posts, nor did they *add*, *follow*, or *friend* any accounts to view their content. RAs used school computers to search, preventing any stored identifiers on their personal devices from inadvertently influencing posts that were shown. This also prevented any bias introduced by algorithmically curated content based on user preferences. RAs copied the posts into Microsoft Excel files for text posts or saved a screen capture of multimedia posts for analysis. RAs added additional search terms based on the content found, and when finding multiple posts from one profile, we then viewed public profiles to capture additional posts. They choose posts for inclusion if they mentioned alcohol (eg, "Fireball helps me study!"), an incident known to be related to alcohol (eg, an assault on a drunk student), pictured alcohol or alcohol containers (eg, photos of discarded red solo cups outside a fraternity house), or implied alcohol use (eg, "Missed class due a late night with the boys!"). RAs assessed each post viewed using search terms published within our timeframe; however, RAs only captured posts that were relevant to alcohol use. For example, we excluded posts such as "The football team was amazing last night!" As underage drinking is an illegal activity, to protect the confidentiality of posters, all posts gathered were public, which limited the number of posts available to the research team. This prompted us to extend the original search period from 5 to 6 years. We also extended the time frame to 6 years to capture reactions to well-known alcohol-related events that occurred on campus or involved students at the university. Each post was screen captured. Images and Excel files containing written social media posts were uploaded into a qualitative analysis software (NVivo 12).

Qualitative Coding Procedures

The data were organized using an inductive grounded theory approach [50]. Magnitude coding was used to count the number of likes received for each post on Facebook, Instagram, Twitter, and YouTube. Categories were created to display the likes each post received. Magnitude coding identified implicit and explicit drinking behavior on Facebook, Instagram, Twitter, and YouTube posts and described students' reactions to posts on social media as positive, negative, or neutral. Open codes included concepts such as the culture surrounding alcohol use at the university, social acceptance, peer pressure, alumni influences, news surrounding tragic alcohol-related events, perceptions of Greek life, and students' ability to drink while maintaining their academic responsibilities.

One researcher reviewed all posts and comments to design a preliminary codebook that defined the codes and described the

coding procedures. Each coder (MNC and JGS) independently coded portions of the data and met weekly to update the codebook with new codes that emerged as more data were coded. An open discussion approach added a third reviewer (JPH) to discuss areas of disagreement [51,52] about how codes were applied to each social media platform. This ensured consistency in how the codes were applied but did not limit the individual coder's view of each post in terms of positivity or explicitness. Coders recoded the data with disagreement, along with the remaining data. After the initial round of independent

coding was completed, we assessed the consistency between the 2 coders using the Cohen κ score. The final κ value was 0.96, reflecting excellent interrater reliability [53]. Twitter was excluded from the overall interrater reliability owing to less than half (809/2183, 37.06%) of the total tweets collected being coded, which resulted in a low overlap of codes. Compared with Instagram, almost all posts were coded as 99% (112/113). Table 1 provides the interrater agreement for the data from each social media platform.

Table 1. Interrater reliability and percent agreement by social media platform.

	Cohen κ score	Percent agreement (%)
Instagram	1.00	100
Facebook	0.78	94.5
YouTube	0.86	96.3
College confidential	0.79	96.2
Greekrank	0.87	97.8
Twitter	0.42	95.6
Reddit	0.35	90.7
Overall interrater reliability	0.96	98.8

Ethics Approval

RAs did not friend, follow, or establish a relationship with anyone who had posted; we only reviewed publicly available posts to avoid violating the privacy or implicit trust of any social media user. This study was approved by the Social and Behavioral Science Institutional Review Board of the University of Virginia (protocol 3282).

Results

Posts by Platform

Multimedia Appendix 1 describes the platforms, how we accessed the platform, why the platform was selected, and the number of posts coded ($n=1151$). Of the social media posts from 2013 to 2019 related to binge drinking at the university, Twitter generated the most posts in this study, followed by Instagram, Greekrank, and Reddit. Fewer posts were gathered from College Confidential, Facebook, and YouTube. Although we considered only using Twitter in our analysis, the other platforms offered a wide variety of viewpoints, attracted users from different perspectives, and reflected posts that may be made at different points in the drinking cycle. Owing to the team's interest in informing intervention development, we included platforms with fewer posts to ensure that data points were from more sources in our analysis.

Social media posts by students before, during, and after alcohol use provided information about students' and peers' real-time attitudes and beliefs. For example, on Instagram, 2 posts illustrated student attitudes toward fake IDs. The first was a video of a creased fake ID containing personally identifiable information. The second showed a photo with the caption "three girls one fake," clearly documenting illegal behavior by multiple students using a fake ID to obtain alcohol while under the legal

age. On YouTube, a video showed a young woman looking for an entrance to the hospital emergency room encountering police, which resulted in her arrest. Many student posts included the phrase "Work hard, play hard." Photos included a picture of a student studying for examinations, surrounded by ≥ 25 open cans of beer. Other posts showed alumni and parents joining in parties that included descriptions or depictions of binge drinking behavior.

Social media posts with positive reactions framed the university as a *party school*, with extreme drinking behavior as the norm. Posts on lifestyle-centered social media platforms, such as Instagram and Greekrank, tended to portray drinking in a glamorous manner, with negative comments reserved for criticizing the quality of parties and events. On platforms where users posted more about the university as an institution of higher learning (Reddit and College Confidential) than about individual student lives, we found more posts that included negative comments about the perceived drinking-centric culture at the school, implying that students who choose not to drink may have a more difficult time fitting in.

Classification of Drinking Attitudes

Drinking attitudes on each platform were classified as positive, neutral, or negative. Table 2 summarizes the attitudes specific to each platform with exemplary posts. Table 3 reports the frequencies of each category of drinking attitudes found on each social media platform. Across all types of posts reviewed, we found that positive drinking attitudes were the most common, followed by negative and then neutral attitudes. These varied by platform. College students' Instagram and Facebook alcohol posts frequently depicted alcohol in a positive social context (118/134, 88.1%), with fewer being classified as neutral (8/134, 6%) or negative (8/134, 6%).

Table 2. Definition of positive, neutral, and negative by platform.

Platform and definition	Example
Instagram	
Positive	The overall tone of the user's attitude is positive and glorifying alcohol consumption making alcohol consumption look glamorous and appealing. <ul style="list-style-type: none"> Two smiling girls sitting on the steps of a deck, obviously under the influence and surrounded by red solo cups and empty bottles, with a caption about what a good day it was.
Neutral	The user is not suggesting an opinion on alcohol. Posts that belong in this category include post about events on campus. <ul style="list-style-type: none"> A photo of a historic marker near several fraternity houses, with alcohol bottles and cups in the background.
Negative	Posts discussing the culture and behaviors of students negatively in regard to over consumption of alcohol. Posts discussing how alcohol is harmful to students and their environment. <ul style="list-style-type: none"> A close-up photo of empty bottles and cups lying in the gutter of a public street with a caption about the unacceptable behavior of university students.
Facebook	
Positive	The overall tone of the user's attitude is positive and glorifying alcohol consumption making alcohol consumption look glamorous and appealing. <ul style="list-style-type: none"> A photo of the membership of a fraternity outside of their house with an American flag and many of the brothers holding drinks in salute.
Neutral	No neutral Facebook posts. N/A ^a
Negative	No negative Facebook posts. N/A
YouTube	
Positive	The overall tone of the user's attitude is positive and glorifying alcohol consumption making alcohol consumption look glamorous and appealing. <ul style="list-style-type: none"> A current student provides a guided tour of the collection of restaurants and bars near campus and advises on which are fun as well as easy to obtain alcohol while underage.
Neutral	No neutral YouTube posts. N/A
Negative	No negative YouTube posts. N/A
Twitter	
Positive	The overall tone of the user's attitude is positive and glorifying alcohol consumption making alcohol consumption look glamorous and appealing. <ul style="list-style-type: none"> "God I love tequila and cute boys who know how to sing!" with hashtags that link it to the university.
Neutral	The user is not suggesting an opinion, just stating a fact or news update. Posts that belong in this category include posts about traumatic events that happened. <ul style="list-style-type: none"> "Alcohol and Drug Abuse Prevention Task Force challenges concept (traditions of excessive drinking by students at specific events)." (includes a link to a newspaper article that prevents several perspectives)
Negative	Posts discussing the culture and behaviors of university students negatively in regard to over consumption of alcohol. Posts discussing how alcohol is harmful to students and their environment. Posts that belong in this category also include posts about user's opinions on traumatic events that happened on campus. <ul style="list-style-type: none"> "(Event) looks like Lily Pulitzer vomited on the entire (school) population." "White privilege, (Event) 2017 style." Both tweets accompanied by photos of students drinking at an event.
College confidential	
Positive	Posts include students talking positively about events on campus and adjusting well into the school culture with and without consumption of alcohol. <ul style="list-style-type: none"> "...but I would go again if I got to do it over (somehow I think you would not go again). A listserv primarily for notifying students of alcohol-free events was updated weekly when I attended; I can't attest to who updates it now or how frequently. I would agree that most people drink even if they are not involved in Greek life ('Thirsty Thursday' is a grounds-wide saying) but I was happy enough sober."

Platform and definition		Example
Neutral	The user is not suggesting an opinion.	<ul style="list-style-type: none"> “Over the past couple of decades, fraternities have gradually been required to use stricter controls on parties. Guest lists, BYOB, etc. This isn’t unique to (this university) though. It’s simply a sign of the times. But it does make it harder for non-Greeks to be part of the Greek social scene. In any event, if joining the Greek system isn’t your thing, it just means you need to build your own social circle whether through clubs and whatnot. Note that after first year, a large number of university students live off campus, so I imagine a lot of Non-Greeks simply have get together at their apartments or even hang out at (local) restaurants and bars.”
Negative	Posts discussing the culture and behaviors of university students negatively in regard to over consumption of alcohol. Students having a hard time adjusting to the culture at the university. Students complaining about excessive alcohol consumption and lack of alternative activities and ways to bond with other students.	<ul style="list-style-type: none"> “Bottom-line, this is NOT a good place to come if you do not plan on being a moderate to heavy party-goer. Yes, you can survive on the periphery of the social scene by not drinking but you will never get that ‘quintessential’ university experience that current students and alum rave about. Don’t let anyone try to convince you otherwise.”
Greekrank		
Positive	Students rating fraternity chapters positively based on their availability of alcohol and social events. Posts in this category include posts that place higher social status on fraternity chapters based on their availability of alcohol, women, and social events.	<ul style="list-style-type: none"> “These guys really know how to party. Except sometimes that causes problems because I pass out at their house, but the guys are super nice and always find a way to get me home back to dorms. Great guys all the way around!” “Top house, biggest parties, coldest beers, hottest women.”
Neutral	Posts in this category include students discussing how fraternity chapters could improve by throwing more social events and having more availability of alcohol. Fraternities in this category are average party goers. These fraternities fall somewhere in the middle between very high social status and very low.	<ul style="list-style-type: none"> “If you’re looking for a mild place to party, good, but if you’re looking for ragers not the place except once a year. Overall ok guys.” “Solid group of guys and definitely a top house but should probably throw more parties to remain a top house.”
Negative	Students rating fraternity chapters negatively based on their lack of availability of alcohol, attractive women, and social events.	<ul style="list-style-type: none"> “A brother puked on me from above in a balcony. They have absolutely no class or alcohol tolerance. Not ‘true Southern gentlemen’ like they think they are.” “Superficial brotherhood. Serves watered down punch.” “Buncha trust fund betas who genuinely think they’re alphas.” “Drink 7 beers and pretend to black out...soft.”
Reddit		
Positive	Posts include students talking positively about events on campus and adjusting well into the university culture with and without consumption of alcohol.	<ul style="list-style-type: none"> “You have to make your own fun. You can try to socialize with people at parties without drinking—go early before people are totally wasted. Fill that solo cup with water and laugh as people get less funny, more incoherent then leave when you’re bored.”
Neutral	The user is not suggesting an opinion, just stating a fact or news update.	<ul style="list-style-type: none"> “You still need an ID at those bars unless you come in early. It doesn’t have to be a good fake but u need one.”
Negative	Posts discussing the culture and behaviors of university students negatively in regard to over consumption of alcohol. Students complaining about excessive alcohol consumption and lack of alternative activities and ways to bond with other students.	<ul style="list-style-type: none"> “It is concerning that the first people encountered at the hospital did nothing to help, other than pointing them to the ER, and that it took so long after the police arrived to begin to move her into the ER—it seems this was dealt with more as a law enforcement matter than an emergency medical situation, which it could have been. That said, it is hard to feel too sorry for either one of the girls, and they should know that people who tend to become angry or combative when drunk have a strong tendency to get themselves in trouble from drinking. I seriously hope they learn something from this.”

^aN/A: not applicable.

Table 3. Attitudes toward drinking by social media platform.

Types of social media	Instagram (n=111 ^a), n (%)	Facebook (n=23), n (%)	Twitter, (n=373 ^a) n (%)	YouTube (n=8), n (%)	Reddit (n=80 ^b), n (%)	College confidential (n=38 ^b), n (%)	Greekrank (n=108 ^b), n (%)
Drinking attitudes							
Positive	95 (85.6)	23 (100)	104 (27.9)	8 (100)	25 (39)	3 (10.7)	48 (44.4)
Negative	8 (7.2)	0 (0)	179 (48)	0 (0)	20 (31.2)	25 (65.8)	47 (43.5)
Neutral	8 (7.2)	0 (0)	90 (24.1)	0 (0)	35 (54.6)	10 (26.3)	13 (12)

^aNot all posts were directly related to drinking attitudes and only those posts that displayed a drinking attitude were included.

^bThis platform included posts that displayed multiple drinking attitudes.

Classification of Alcohol Use Depictions

Drinking experiences were classified as implicit or explicit on platforms with photographs. Explicit alcohol use was common on Instagram and Facebook, seen in 66.7% (74/111) of Instagram posts, whereas the remaining Instagram posts alluded to alcohol consumption and were coded as implicit (37/111, 33.3%). Facebook posts were usually explicit in depicting alcohol use (21/23, 91%), with a few implicit consumption posts (2/23, 9%). Facebook, Instagram, and YouTube used a visual medium and were grouped together for the analysis. Posts on these platforms (with similar definitions of positive, neutral, and negative) showed more explicit alcohol consumption (102/140, 72.9%; Table 2). Nearly all posts containing explicit alcohol consumption were depicted as positive by the person posting (99/102, 97%); only 3% (3/102) were portrayed as negative. Similarly, implicit alcohol consumption was mostly portrayed as positive (31/38, 81%), with only 18% (7/38) of these posts seemingly negative by the person posting. All YouTube posts were categorized as positive and explicit.

In contrast, posts on Reddit, College Confidential, and Greekrank showed a wider range of views on alcohol events. Overall, 31% (20/64) of the Reddit posts portrayed alcohol consumption as negative, 54% (35/64) as neutral, and 39% (25/64) as positive. College Confidential frequently portrayed alcohol consumption as negative (25/34, 65%), with a minority of posts about alcohol (3/34, 10%) being positive. Greekrank

posts were evenly split as positive (48/71, 44%) or negative (47/71, 43%). Compared with Instagram, Facebook, and YouTube, College Confidential, Greekrank, and Reddit had fewer alcohol-related posts depicted as positive. Finally, we identified high-risk drinking events in the local areas where college students tend to frequently binge drink. Examples included gatherings at local bars and restaurants, festivals, fraternity houses, yearly events, and traditions around academic breaks, seasons of the year, and sports events. Most of these events were displayed positively. Negative events (eg, student injuries) were positive toward the student but negative toward the university administration. Of the posts that exhibited high-risk drinking events, most (274/370, 74.1%) were shown in a positive social context, whereas only 25.9% (96/370) were shown in a negative social context.

Peer Reactions to Posts

On average, Instagram posts received 350 *likes*, or positive feedback tags placed on posts by viewers other than the original poster to indicate their reaction (mean 350.1, SD 1224.38). More than half of all Instagram posts had over 100 likes (65/111, 58.6%). On Instagram, only 12.6% (14/111) of the posts had fewer than 20 likes. Facebook posts received 43 likes on average (mean 43.3, SD 51.91), and only 17% (4/23) of Facebook posts had over 100 likes, whereas nearly half of all Facebook posts had fewer than 20 likes for each post (11/23, 48%). Table 4 summarizes peer feedback on these posts.

Table 4. Implicit and explicit alcohol content on Instagram, Facebook, and YouTube.

Drinking behaviors	Instagram (n=111), n (%)	Facebook (n=23), n (%)	YouTube (n=8), n (%)
Implicit alcohol content	37 (33.3)	2 (8.7)	0 (0)
Explicit alcohol content	74 (66.7)	21 (91.3)	8 (100)
Count of "likes" (if videos, count of "views")			
<20	14 (12.6)	11 (47.8)	0 (0)
20-50	15 (13.5)	6 (26.1)	0 (0)
50-100	17 (15.3)	2 (8.7)	1 (12.5)
100-500	51 (45.9)	4 (17.4)	1 (12.5)
≥500	14 (12.6)	0 (0)	6 (75)

Common Themes

As we worked with the data, the coders noticed that certain words frequently appeared. The largest words showed the highest frequency, included alcohol, culture, Greek, drunk, and

party. The visual representation of the word cloud shows the common words in the content of social media posts.

Discussion

Principal Findings

Students post alcohol use content, which was then liked, retweeted, and shared, likely reinforcing the poster, who is creating a public image of a desirable life to their peers. Students freely shared their drinking behaviors and attitudes toward alcohol use on publicly accessible social media platforms [54,55], visible not only by peers but also by school administrators, parents, and future employers. Thus, the impact of student posts could extend beyond current social networks, but posts reflected little concern about how social media history could influence future prospects.

Alcohol-related events including student deaths have made headlines. Some social media posts showed that students believed that the university should have done more to prevent these tragedies. In other cases, social media posts indicated an attitude of blaming the victims for being in these dangerous situations. Few posts mentioned the dangers related to excessive drinking or taking responsibility for drinking behavior. Instead, more posts showed students in situations where they could use help and guidance, such as posts showing students passed out in public spaces, intoxicated while walking on the streets, and other situations where their safety and well-being are at risk. These posts demonstrate that students may need a convenient, easy-to-understand resource that explains the university's policies on alcohol use and how to obtain help. This guidance should be available at the moment of excessive drinking.

Several social media posts included multiple generations of alumni and current students drinking alcohol together. These posts demonstrate that drinking is normative and multigenerational. This implies that students must drink alcohol to have a typical, happy experience at the university. Although there were some posts critical of the portrayed drinking culture, a few posts showed that moderate drinking is an option. A few posts shared the idea that it is possible to have a satisfying, fun, and enjoyable college social life without participating in binge drinking. Some students indicated a desire for greater availability of events and opportunities that are not centered on alcohol use, both by students who abstain and those who would like a break from always drinking when socializing.

Limitations

We did not capture every social media post on the selected platforms nor did we capture data from all possible social media platforms. For example, Snapchat is a popular platform in which posts are time limited, with privacy protections that limit the public availability of posts, making it unavailable for data collection. Snapchat videos, also known as *stories*, were not investigated, as they cannot be seen unless you are a friend of the poster, which adds a limitation. Stories are deleted after 24 hours, and it is possible that it is a key platform for sharing binge drinking [54] that we were unable to access. Platforms such as Facebook and Instagram have increased privacy options, and users may restrict nonfollowers from viewing their content. By limiting the searches for this study to publicly available posts, we may have missed content that portrays different drinking norms. In addition, searching platforms such as Twitter

is complicated, as hashtag searches often yield numerous posts that are not relevant to the topic. For example, searching for the university name resulted in posts about sports teams, faculty news, and other non-student-generated content. Despite these limitations, we captured a representative sample of public student posts from different platforms. The student perspectives expressed on publicly available social media show little evidence of *faking good* that could occur when participants are interviewed or surveyed.

Comparison With Prior Work

Other studies on social media and binge drinking have examined responses to specific messages [56,57], the number of posts [58], and habits of social media use [58,59] compared with reported drinking behavior. This study adds to the literature by analyzing the content of posts and responses to posts from multiple platforms. These data create a picture of drinking behaviors at the university that has implications for the development of interventions. Specifically, this study identified 5 concerning themes that the interventions could address. First, public posts clearly showed students engaging in risky drinking behaviors. Tips and tools for managing drinking (and posting about it) in real time are needed. Second, several posts showed that students who did not know how to access or ask for help because of concerns that they or their friends would face consequences. Third, social media posts depict a norm that all students at the university drink heavily and that a happy social life depends upon binge drinking. Fourth, these public posts could sometimes harm student reputation. Fifth, several posts showed that parents and alumni were part of the drinking culture, including posts made by parents while drinking with students.

On the basis of these themes, we offer 5 recommendations to reduce harm related to excessive drinking among college students. First, the early timing (often before or during the first semester of college) and universal targets of alcohol education should be reconsidered. In addition to existing alcohol education, we assert that there is a need for tailored, easy-to-use tools that students can use in real time when drinking. Second, universities should show students how to obtain help from themselves or their peers during drinking without penalty, even if they are underage. Third, students at the university (and likely others) need assistance in finding popular alcohol-free social activities. Fourth, colleges should raise students' awareness of their web-based reputations and provide options to help them repair their web-based reputations if public posts show them under the influence of alcohol. Fifth, universities should encourage parents and alumni to modify their own drinking habits at university events or gatherings to provide better role modeling. Previous studies have found that interventions encouraging parents to model acceptable limits for alcohol consumption can have a positive impact on delaying and reducing student drinking [60-64].

Conclusions

This qualitative analysis of social media posts on college student drinking is the first to characterize student posting and commenting behavior across multiple social media platforms. This adds to a growing body of literature showing that analyzing social media can reveal the context of hazardous drinking

behavior [65]. An important contribution of this study is that it demonstrated the attitudes and actions of students during binge drinking; these may differ substantially from what students report on alcohol use surveys. Important and novel findings are as follows: (1) social media platforms are being used before, during, and after the time of hazardous drinking; (2) most posts showing consequences of excessive drinking occurred in near real time; (3) the majority of posts showing explicit alcohol consumption were positively reciprocated with more likes and comments indicating students' positive attitudes toward risky drinking behaviors; peers' comments about their peers' posts create a web-based social context that strongly reinforces risky drinking behavior; and (4) students are often depicted in risky situations when drinking, and posts about these could damage student reputations or future prospects. The study identified students' thoughts and beliefs about the binge drinking culture

at one university, but it is likely that the concerning themes and resulting recommendations will be generalized to other colleges.

Finally, this social media review identifies several new targets for intervention. Students need real-time interventions during their drinking episodes before they experience harm. Students lack awareness of resources for improving safety for themselves and their peers while drinking and need to access this guidance when witnessing excess drinking. Students also need consciousness-raising interventions regarding the risks of creating or allowing social media posts during drinking episodes. These and other behavioral targets could be addressed by platforms such as mobile apps that could provide information during drinking episodes. A thoughtfully developed mobile app could provide tailored, real-time tracking of drinking behavior; guidance to improve student drinking safety; and reminders against posting content that could harm reputations and limit future options.

Acknowledgments

The study was funded by the University of Virginia Strategic Investment Fund under the program "Reducing the Burden of Addiction in Virginia." NSA is now at George Mason University in Fairfax, Virginia. MNC is currently at the Recovery Research Institute in Boston, Massachusetts.

Conflicts of Interest

None declared.

Multimedia Appendix 1

Social media platform descriptions, search methods, restrictions, and yield for the study.

[[DOCX File, 18 KB - humanfactors_v9i2e36239_app1.docx](#)]

References

1. Johnston LD, O'Malley PM, Bachman JG. Monitoring the future -- national results on adolescent drug use: overview of key findings 2004. Office of Justice Programs, U.S. Department of Justice. 2005 Apr. URL: <https://www.ojp.gov/ncjrs/virtual-library/abstracts/monitoring-future-national-results-adolescent-drug-use-overview-3> [accessed 2022-05-12]
2. College Drinking. National Institute on Alcohol Abuse and Alcoholism. 2021 Oct. URL: <https://www.niaaa.nih.gov/publications/brochures-and-fact-sheets/college-drinking> [accessed 2021-08-09]
3. Wechsler H, Dowdall GW, Maenner G, Gledhill-Hoyt J, Lee H. Changes in binge drinking and related problems among American college students between 1993 and 1997. Results of the Harvard School of Public Health College Alcohol Study. *J Am Coll Health* 1998 Sep;47(2):57-68. [doi: [10.1080/07448489809595621](https://doi.org/10.1080/07448489809595621)] [Medline: [9782661](https://pubmed.ncbi.nlm.nih.gov/9782661/)]
4. American College Health Association-National College Health Assessment II: Reference Group Executive Summary. American College Health Association. 2019. URL: https://www.acha.org/documents/ncha/NCHA-II_SPRING_2019_US_REFERENCE_GROUP_EXECUTIVE_SUMMARY.pdf [accessed 2021-10-06]
5. Hingson R, Zha W, Smyth D. Magnitude and trends in heavy episodic drinking, alcohol-impaired driving, and alcohol-related mortality and overdose hospitalizations among emerging adults of college ages 18-24 in the United States, 1998-2014. *J Stud Alcohol Drugs* 2017 Jul;78(4):540-548 [FREE Full text] [doi: [10.15288/jsad.2017.78.540](https://doi.org/10.15288/jsad.2017.78.540)] [Medline: [28728636](https://pubmed.ncbi.nlm.nih.gov/28728636/)]
6. 2015 National Survey of Drug Use and Health (NSDUH) Releases. Substance Abuse and Mental Health Services Administration. 2015. URL: <https://www.samhsa.gov/data/release/2015-national-survey-drug-use-and-health-nsduh-releases> [accessed 2021-10-06]
7. McCabe SE, Veliz P, Schulenberg JE. How collegiate fraternity and sorority involvement relates to substance use during young adulthood and substance use disorders in early midlife: a national longitudinal study. *J Adolesc Health* 2018 Mar;62(3S):S35-S43 [FREE Full text] [doi: [10.1016/j.jadohealth.2017.09.029](https://doi.org/10.1016/j.jadohealth.2017.09.029)] [Medline: [29455716](https://pubmed.ncbi.nlm.nih.gov/29455716/)]
8. Kuntsche E, Kuntsche S, Thrul J, Gmel G. Binge drinking: health impact, prevalence, correlates and interventions. *Psychol Health* 2017 Aug;32(8):976-1017. [doi: [10.1080/08870446.2017.1325889](https://doi.org/10.1080/08870446.2017.1325889)] [Medline: [28513195](https://pubmed.ncbi.nlm.nih.gov/28513195/)]
9. Rinker DV, Krieger H, Neighbors C. Social network factors and addictive behaviors among college students. *Curr Addict Rep* 2016 Dec;3(4):356-367 [FREE Full text] [doi: [10.1007/s40429-016-0126-7](https://doi.org/10.1007/s40429-016-0126-7)] [Medline: [28580226](https://pubmed.ncbi.nlm.nih.gov/28580226/)]

10. Steers ML, Moreno MA, Neighbors C. The influence of social media on addictive behaviors in college students. *Curr Addict Rep* 2016 Dec;3(4):343-348 [FREE Full text] [doi: [10.1007/s40429-016-0123-x](https://doi.org/10.1007/s40429-016-0123-x)] [Medline: [28458990](https://pubmed.ncbi.nlm.nih.gov/28458990/)]
11. Young SD. Behavioral insights on big data: using social media for predicting biomedical outcomes. *Trends Microbiol* 2014 Nov;22(11):601-602 [FREE Full text] [doi: [10.1016/j.tim.2014.08.004](https://doi.org/10.1016/j.tim.2014.08.004)] [Medline: [25438614](https://pubmed.ncbi.nlm.nih.gov/25438614/)]
12. Gruebner O, Sykora M, Lowe SR, Shankardass K, Galea S, Subramanian SV. Big data opportunities for social behavioral and mental health research. *Soc Sci Med* 2017 Sep;189:167-169. [doi: [10.1016/j.socscimed.2017.07.018](https://doi.org/10.1016/j.socscimed.2017.07.018)] [Medline: [28755794](https://pubmed.ncbi.nlm.nih.gov/28755794/)]
13. Culotta A. Lightweight methods to estimate influenza rates and alcohol sales volume from Twitter messages. *Lang Resour Eval* 2013;47(1):217-238. [doi: [10.1007/s10579-012-9185-0](https://doi.org/10.1007/s10579-012-9185-0)]
14. Culotta A. Estimating county health statistics with Twitter. In: *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. 2014 Presented at: CHI '14; April 26-May 1, 2014; Toronto, Canada p. 1335-1344. [doi: [10.1145/2556288.2557139](https://doi.org/10.1145/2556288.2557139)]
15. Eichstaedt JC, Schwartz HA, Kern ML, Park G, Labarthe DR, Merchant RM, et al. Psychological language on Twitter predicts county-level heart disease mortality. *Psychol Sci* 2015 Feb;26(2):159-169 [FREE Full text] [doi: [10.1177/0956797614557867](https://doi.org/10.1177/0956797614557867)] [Medline: [25605707](https://pubmed.ncbi.nlm.nih.gov/25605707/)]
16. Ginsberg J, Mohebbi MH, Patel RS, Brammer L, Smolinski MS, Brilliant L. Detecting influenza epidemics using search engine query data. *Nature* 2009 Feb 19;457(7232):1012-1014. [doi: [10.1038/nature07634](https://doi.org/10.1038/nature07634)] [Medline: [19020500](https://pubmed.ncbi.nlm.nih.gov/19020500/)]
17. Jena AB, Karaca-Mandic P, Weaver L, Seabury SA. Predicting new diagnoses of HIV infection using Internet search engine data. *Clin Infect Dis* 2013 May;56(9):1352-1353. [doi: [10.1093/cid/cit022](https://doi.org/10.1093/cid/cit022)] [Medline: [23334812](https://pubmed.ncbi.nlm.nih.gov/23334812/)]
18. Paul M, Dredze M. You are what you Tweet: analyzing Twitter for public health. In: *Proceedings of the 5th International AAAI Conference on Web and Social Media*. 2011 Presented at: ICWSM '11; July 17-21, 2011; Barcelona, Spain p. 265-272.
19. Most popular social networks of teenagers in the United States from fall 2012 to fall 2021. Statista. URL: <https://www.statista.com/statistics/250172/social-network-usage-of-us-teens-and-young-adults/> [accessed 2020-08-14]
20. Internet and technology: mobile fact sheet. Pew Research Center. 2021 Apr 7. URL: <https://www.pewresearch.org/internet/fact-sheet/mobile/> [accessed 2021-11-30]
21. Moreno MA, Whitehill JM. Influence of social media on alcohol use in adolescents and young adults. *Alcohol Res* 2014;36(1):91-100 [FREE Full text] [Medline: [26259003](https://pubmed.ncbi.nlm.nih.gov/26259003/)]
22. Hendriks H, Van den Putte B, Gebhardt WA, Moreno MA. Social drinking on social media: content analysis of the social aspects of alcohol-related posts on Facebook and Instagram. *J Med Internet Res* 2018 Jun 22;20(6):e226 [FREE Full text] [doi: [10.2196/jmir.9355](https://doi.org/10.2196/jmir.9355)] [Medline: [29934290](https://pubmed.ncbi.nlm.nih.gov/29934290/)]
23. Griffiths R, Casswell S. Intoxicogenic digital spaces? Youth, social networking sites and alcohol marketing. *Drug Alcohol Rev* 2010 Sep;29(5):525-530. [doi: [10.1111/j.1465-3362.2010.00178.x](https://doi.org/10.1111/j.1465-3362.2010.00178.x)] [Medline: [20887576](https://pubmed.ncbi.nlm.nih.gov/20887576/)]
24. Moreno MA, Arseniev-Koehler A, Litt D, Christakis D. Evaluating college students' displayed alcohol references on Facebook and Twitter. *J Adolesc Health* 2016 May;58(5):527-532 [FREE Full text] [doi: [10.1016/j.jadohealth.2016.01.005](https://doi.org/10.1016/j.jadohealth.2016.01.005)] [Medline: [26995291](https://pubmed.ncbi.nlm.nih.gov/26995291/)]
25. Boyle SC, LaBrie JW, Froidevaux NM, Witkovic YD. Different digital paths to the keg? How exposure to peers' alcohol-related social media content influences drinking among male and female first-year college students. *Addict Behav* 2016 Jun;57:21-29 [FREE Full text] [doi: [10.1016/j.addbeh.2016.01.011](https://doi.org/10.1016/j.addbeh.2016.01.011)] [Medline: [26835604](https://pubmed.ncbi.nlm.nih.gov/26835604/)]
26. van Hoof JJ, Bekkers J, van Vuuren MA. Son, you're smoking on Facebook! College students' disclosures on social networking sites as indicators of real-life risk behaviors. *Comput Human Behav* 2014 May;34:249-257. [doi: [10.1016/j.chb.2014.02.008](https://doi.org/10.1016/j.chb.2014.02.008)]
27. Giorgi S, Yaden DB, Eichstaedt JC, Ashford RD, Buffone AE, Schwartz HA, et al. Cultural differences in Tweeting about drinking across the US. *Int J Environ Res Public Health* 2020 Feb 11;17(4):1125 [FREE Full text] [doi: [10.3390/ijerph17041125](https://doi.org/10.3390/ijerph17041125)] [Medline: [32053866](https://pubmed.ncbi.nlm.nih.gov/32053866/)]
28. Smith RL, Salvatore JE, Aliev F, Neale Z, Barr P, Spit for Science Working Group, et al. Genes, roommates, and residence halls: a multidimensional study of the role of peer drinking on college students' alcohol use. *Alcohol Clin Exp Res* 2019 Jun;43(6):1254-1262. [doi: [10.1111/acer.14037](https://doi.org/10.1111/acer.14037)] [Medline: [31034622](https://pubmed.ncbi.nlm.nih.gov/31034622/)]
29. Hoffman EW, Austin EW, Pinkleton BE, Austin BW. An exploration of the associations of alcohol-related social media use and message interpretation outcomes to problem drinking among college students. *Health Commun* 2017 Jul;32(7):864-871. [doi: [10.1080/10410236.2016.1195677](https://doi.org/10.1080/10410236.2016.1195677)] [Medline: [27421038](https://pubmed.ncbi.nlm.nih.gov/27421038/)]
30. Huang GC, Unger JB, Soto D, Fujimoto K, Pentz MA, Jordan-Marsh M, et al. Peer influences: the impact of online and offline friendship networks on adolescent smoking and alcohol use. *J Adolesc Health* 2014 May;54(5):508-514 [FREE Full text] [doi: [10.1016/j.jadohealth.2013.07.001](https://doi.org/10.1016/j.jadohealth.2013.07.001)] [Medline: [24012065](https://pubmed.ncbi.nlm.nih.gov/24012065/)]
31. Geusens F, Beullens K. The reciprocal associations between sharing alcohol references on social networking sites and binge drinking: a longitudinal study among late adolescents. *Comput Human Behav* 2017 Aug;73:499-506. [doi: [10.1016/j.chb.2017.03.062](https://doi.org/10.1016/j.chb.2017.03.062)]
32. Litt DM, Stock ML. Adolescent alcohol-related risk cognitions: the roles of social norms and social networking sites. *Psychol Addict Behav* 2011 Dec;25(4):708-713. [doi: [10.1037/a0024226](https://doi.org/10.1037/a0024226)] [Medline: [21644803](https://pubmed.ncbi.nlm.nih.gov/21644803/)]

33. Nesi J, Rothenberg WA, Hussong AM, Jackson KM. Friends' alcohol-related social networking site activity predicts escalations in adolescent drinking: mediation by peer norms. *J Adolesc Health* 2017 Jun;60(6):641-647 [FREE Full text] [doi: [10.1016/j.jadohealth.2017.01.009](https://doi.org/10.1016/j.jadohealth.2017.01.009)] [Medline: [28325545](https://pubmed.ncbi.nlm.nih.gov/28325545/)]
34. Tucker JS, Miles JN, D'Amico EJ. Cross-lagged associations between substance use-related media exposure and alcohol use during middle school. *J Adolesc Health* 2013 Oct;53(4):460-464 [FREE Full text] [doi: [10.1016/j.jadohealth.2013.05.005](https://doi.org/10.1016/j.jadohealth.2013.05.005)] [Medline: [23770074](https://pubmed.ncbi.nlm.nih.gov/23770074/)]
35. Cavazos-Rehg PA, Krauss MJ, Sowles SJ, Bierut LJ. "Hey everyone, I'm drunk." An evaluation of drinking-related Twitter chatter. *J Stud Alcohol Drugs* 2015 Jul;76(4):635-643 [FREE Full text] [doi: [10.15288/jsad.2015.76.635](https://doi.org/10.15288/jsad.2015.76.635)] [Medline: [26098041](https://pubmed.ncbi.nlm.nih.gov/26098041/)]
36. Thompson L, Rivara FP, Whitehill JM. Prevalence of marijuana-related traffic on Twitter, 2012-2013: a content analysis. *Cyberpsychol Behav Soc Netw* 2015 Jun;18(6):311-319 [FREE Full text] [doi: [10.1089/cyber.2014.0620](https://doi.org/10.1089/cyber.2014.0620)] [Medline: [26075917](https://pubmed.ncbi.nlm.nih.gov/26075917/)]
37. Beullens K, Schepers A. Display of alcohol use on Facebook: a content analysis. *Cyberpsychol Behav Soc Netw* 2013 Jul;16(7):497-503. [doi: [10.1089/cyber.2013.0044](https://doi.org/10.1089/cyber.2013.0044)] [Medline: [23617225](https://pubmed.ncbi.nlm.nih.gov/23617225/)]
38. Lyons AC, Goodwin I, McCreanor T, Griffin C. Social networking and young adults' drinking practices: innovative qualitative methods for health behavior research. *Health Psychol* 2015 Apr;34(4):293-302. [doi: [10.1037/hea0000168](https://doi.org/10.1037/hea0000168)] [Medline: [25822049](https://pubmed.ncbi.nlm.nih.gov/25822049/)]
39. Moreno MA, D'Angelo J, Kacvinsky LE, Kerr B, Zhang C, Eickhoff J. Emergence and predictors of alcohol reference displays on Facebook during the first year of college. *Comput Human Behav* 2014 Jan;30:10 [FREE Full text] [doi: [10.1016/j.chb.2013.07.060](https://doi.org/10.1016/j.chb.2013.07.060)] [Medline: [24415846](https://pubmed.ncbi.nlm.nih.gov/24415846/)]
40. Moreno MA, Cox ED, Young HN, Haaland W. Underage college students' alcohol displays on Facebook and real-time alcohol behaviors. *J Adolesc Health* 2015 Jun;56(6):646-651 [FREE Full text] [doi: [10.1016/j.jadohealth.2015.02.020](https://doi.org/10.1016/j.jadohealth.2015.02.020)] [Medline: [26003580](https://pubmed.ncbi.nlm.nih.gov/26003580/)]
41. Cabrera-Nguyen EP, Cavazos-Rehg P, Krauss M, Bierut LJ, Moreno MA. Young adults' exposure to alcohol- and marijuana-related content on Twitter. *J Stud Alcohol Drugs* 2016 Mar;77(2):349-353 [FREE Full text] [doi: [10.15288/jsad.2016.77.349](https://doi.org/10.15288/jsad.2016.77.349)] [Medline: [26997194](https://pubmed.ncbi.nlm.nih.gov/26997194/)]
42. Westgate EC, Neighbors C, Heppner H, Jahn S, Lindgren KP. "I will take a shot for every 'like' I get on this status": posting alcohol-related Facebook content is linked to drinking outcomes. *J Stud Alcohol Drugs* 2014 May;75(3):390-398 [FREE Full text] [doi: [10.15288/jsad.2014.75.390](https://doi.org/10.15288/jsad.2014.75.390)] [Medline: [24766750](https://pubmed.ncbi.nlm.nih.gov/24766750/)]
43. D'Angelo J, Kerr B, Moreno MA. Facebook displays as predictors of binge drinking: from the virtual to the visceral. *Bull Sci Technol Soc* 2014;34(5-6):159-169 [FREE Full text] [doi: [10.1177/0270467615584044](https://doi.org/10.1177/0270467615584044)] [Medline: [26412923](https://pubmed.ncbi.nlm.nih.gov/26412923/)]
44. Ridout B, Campbell A, Ellis L. 'Off your Face(book)': alcohol in online social identity construction and its relation to problem drinking in university students. *Drug Alcohol Rev* 2012 Jan;31(1):20-26. [doi: [10.1111/j.1465-3362.2010.00277.x](https://doi.org/10.1111/j.1465-3362.2010.00277.x)] [Medline: [21355935](https://pubmed.ncbi.nlm.nih.gov/21355935/)]
45. Cline S. Playboy: UVA Is Nation's Top Party School. *US News*. 2012 Sep 26. URL: <https://www.usnews.com/news/articles/2012/09/26/playboy-uva-is-nations-top-party-school-playboy-uva-is-nations-top-party-school> [accessed 2021-10-27]
46. 2022 Top Party Schools in America. *Niche*. URL: <https://www.niche.com/colleges/search/top-party-schools/?page=2> [accessed 2021-10-27]
47. Abrams Z. The top 50 party schools in America. *Stacker*. 2019 Jun 28. URL: <https://stacker.com/stories/3217/top-50-party-schools-america> [accessed 2021-10-27]
48. Top Party Schools. *School Authority*. URL: <https://schoolauthority.org/l/top-party-schools> [accessed 2022-05-12]
49. Hsieh HF, Shannon SE. Three approaches to qualitative content analysis. *Qual Health Res* 2005 Nov;15(9):1277-1288. [doi: [10.1177/1049732305276687](https://doi.org/10.1177/1049732305276687)] [Medline: [16204405](https://pubmed.ncbi.nlm.nih.gov/16204405/)]
50. Chapman AL, Hadfield M, Chapman CJ. Qualitative research in healthcare: an introduction to grounded theory using thematic analysis. *J R Coll Physicians Edinb* 2015;45(3):201-205. [doi: [10.4997/JRCPE.2015.305](https://doi.org/10.4997/JRCPE.2015.305)] [Medline: [26517098](https://pubmed.ncbi.nlm.nih.gov/26517098/)]
51. Chan GJ, Getnet M, Olowojesiku R, Min-Swe T, Hunegnaw B, Bekele D. Scoping review protocol on maternal, newborn and child health research in Ethiopia. *BMJ Open* 2020 Jul 02;10(7):e034307 [FREE Full text] [doi: [10.1136/bmjopen-2019-034307](https://doi.org/10.1136/bmjopen-2019-034307)] [Medline: [32616486](https://pubmed.ncbi.nlm.nih.gov/32616486/)]
52. Roth S. Temple University Libraries. 2022 Mar 29. URL: <https://guides.temple.edu/c.php?g=78618&p=3879604> [accessed 2022-04-12]
53. McHugh ML. Interrater reliability: the kappa statistic. *Biochem Med (Zagreb)* 2012;22(3):276-282 [FREE Full text] [Medline: [23092060](https://pubmed.ncbi.nlm.nih.gov/23092060/)]
54. Vranken S, Geusens F, Meeus A, Beullens K. The platform is the message? Exploring the relation between different social networking sites and different forms of alcohol use. *Health New Med Res* 2020;4(2):135-168 [FREE Full text]
55. Bartley N, Abeliuk A, Ferrara E, Lerman K. Auditing algorithmic bias on Twitter. In: *Proceedings of the 13th ACM Web Science Conference*. 2021 Presented at: WebSci '21; June 21-25, 2001; Virtual p. 65-73. [doi: [10.1145/3447535.3462491](https://doi.org/10.1145/3447535.3462491)]
56. Kessler TA, Kurtz CP. Influencing the binge-drinking culture on a college campus. *Nurse Educ* 2019;44(2):106-111. [doi: [10.1097/NNE.0000000000000559](https://doi.org/10.1097/NNE.0000000000000559)] [Medline: [29994998](https://pubmed.ncbi.nlm.nih.gov/29994998/)]

57. Yang B, Zhao X. TV, social media, and college students' binge drinking intentions: moderated mediation models. *J Health Commun* 2018;23(1):61-71. [doi: [10.1080/10810730.2017.1411995](https://doi.org/10.1080/10810730.2017.1411995)] [Medline: [29265924](https://pubmed.ncbi.nlm.nih.gov/29265924/)]
58. Brunborg GS, Andreas JB, Kvaavik E. Social media use and episodic heavy drinking among adolescents. *Psychol Rep* 2017 Jun;120(3):475-490. [doi: [10.1177/0033294117697090](https://doi.org/10.1177/0033294117697090)] [Medline: [28558617](https://pubmed.ncbi.nlm.nih.gov/28558617/)]
59. Ceballos NA, Howard K, Dailey S, Sharma S, Grimes T. Collegiate binge drinking and social media use among Hispanics and non-Hispanics. *J Stud Alcohol Drugs* 2018 Nov;79(6):868-875. [doi: [10.15288/jsad.2018.79.868](https://doi.org/10.15288/jsad.2018.79.868)] [Medline: [30573017](https://pubmed.ncbi.nlm.nih.gov/30573017/)]
60. Parenting to Prevent Childhood Alcohol Use. National Institute on Alcohol Abuse and Alcoholism. 2021 May. URL: <https://www.niaaa.nih.gov/publications/brochures-and-fact-sheets/parenting-prevent-childhood-alcohol-use> [accessed 2021-10-06]
61. Carver H, Elliott L, Kennedy C, Hanley J. Parent-child connectedness and communication in relation to alcohol, tobacco and drug use in adolescence: an integrative review of the literature. *Drugs Educ Prev Policy* 2017;24(2):119-133. [doi: [10.1080/09687637.2016.1221060](https://doi.org/10.1080/09687637.2016.1221060)]
62. Varvil-Weld L, Scaglione N, Cleveland MJ, Mallett KA, Turrisi R, Abar CC. Optimizing timing and dosage: does parent type moderate the effects of variations of a parent-based intervention to reduce college student drinking? *Prev Sci* 2014 Feb;15(1):94-102 [FREE Full text] [doi: [10.1007/s11121-012-0356-4](https://doi.org/10.1007/s11121-012-0356-4)] [Medline: [23404668](https://pubmed.ncbi.nlm.nih.gov/23404668/)]
63. Dumas DM, Turrisi R, Ray AE, Esp SM, Curtis-Schaeffer AK. A randomized trial evaluating a parent based intervention to reduce college drinking. *J Subst Abuse Treat* 2013 Jul;45(1):31-37 [FREE Full text] [doi: [10.1016/j.jsat.2012.12.008](https://doi.org/10.1016/j.jsat.2012.12.008)] [Medline: [23369415](https://pubmed.ncbi.nlm.nih.gov/23369415/)]
64. Ichiyama MA, Fairlie AM, Wood MD, Turrisi R, Francis DP, Ray AE, et al. A randomized trial of a parent-based intervention on drinking behavior among incoming college freshmen. *J Stud Alcohol Drugs Suppl* 2009 Jul(16):67-76 [FREE Full text] [doi: [10.15288/jsads.2009.s16.67](https://doi.org/10.15288/jsads.2009.s16.67)] [Medline: [19538914](https://pubmed.ncbi.nlm.nih.gov/19538914/)]
65. Knox J, Schneider J, Greene E, Nicholson J, Hasin D, Sandfort T. Using social network analysis to examine alcohol use among adults: a systematic review. *PLoS One* 2019 Aug 22;14(8):e0221360 [FREE Full text] [doi: [10.1371/journal.pone.0221360](https://doi.org/10.1371/journal.pone.0221360)] [Medline: [31437257](https://pubmed.ncbi.nlm.nih.gov/31437257/)]

Abbreviations

RA: research assistant

Edited by A Kushniruk; submitted 06.01.22; peer-reviewed by H Stevens, S Moore; comments to author 21.02.22; revised version received 16.04.22; accepted 19.04.22; published 30.05.22.

Please cite as:

Cirillo MN, Halbert JP, Smith JG, Alamiri NS, Ingersoll KS

#BingeDrinking—Using Social Media to Understand College Binge Drinking: Qualitative Study

JMIR Hum Factors 2022;9(2):e36239

URL: <https://humanfactors.jmir.org/2022/2/e36239>

doi: [10.2196/36239](https://doi.org/10.2196/36239)

PMID: [35635740](https://pubmed.ncbi.nlm.nih.gov/35635740/)

©Madison N Cirillo, Jennifer P Halbert, Jessica Gomez Smith, Nour Sami Alamiri, Karen S Ingersoll. Originally published in *JMIR Human Factors* (<https://humanfactors.jmir.org>), 30.05.2022. This is an open-access article distributed under the terms of the Creative Commons Attribution License (<https://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in *JMIR Human Factors*, is properly cited. The complete bibliographic information, a link to the original publication on <https://humanfactors.jmir.org>, as well as this copyright and license information must be included.

Original Paper

Features and Components Preferred by Adolescents in Smartphone Apps for the Promotion of Physical Activity: Focus Group Study

Alex Domin¹, BA, MPhil, PhD; Yacine Ouzzahra², BSc, MSc, PhD; Claus Vögele¹, BSc, MSc, PhD

¹Research Group for Self-Regulation and Health, Department of Behavioural and Cognitive Sciences, University of Luxembourg, Esch-sur-Alzette, Luxembourg

²Research Support Department, University of Luxembourg, Esch-sur-Alzette, Luxembourg

Corresponding Author:

Alex Domin, BA, MPhil, PhD

Research Group for Self-Regulation and Health

Department of Behavioural and Cognitive Sciences

University of Luxembourg

Maison des Sciences Humaines

11, Porte des Sciences

Esch-sur-Alzette, L-4366

Luxembourg

Phone: 352 46 66 44 9389

Fax: 352 46 66 44 39389

Email: alex.domin@uni.lu

Abstract

Background: There is solid evidence that lack of physical activity (PA) is a risk factor for chronic diseases. Sufficient levels of PA in childhood and adolescence are particularly important, as they can set the standards for PA levels in adulthood. The latest reports show that only a small percentage of adolescents reach the recommended levels of PA in European Union countries at the age of 15 years. In view of the scale of the problem, it is crucial to develop interventions that promote and support PA in adolescents. Considering their low implementation costs and ubiquitous presence, smartphone apps could be advantageous as a part of PA interventions.

Objective: This study aimed at investigating the attitudes and preferences of adolescents aged 16-18 years toward various PA app features and components that could (1) make the app more attractive for them and consequently (2) increase their interest and engagement with the app.

Methods: Two separate focus group discussions were conducted in 2 groups of adolescents (n=4 each) aged 16-18 years. Focus groups were carried out online via video conference. The discussions were conducted using a semistructured interview. Participants (n=8; 4 males and 4 females) had a mean age of 17.25 years (SD 0.82 years). Transcripts were analyzed following the approach by Krueger and Casey, that is, categorizing participants' answers and comments according to the questions and themes from the focus group schedule.

Results: Features, such as "goal setting and planning," "coaching and training programs," "activity tracking," "feedback," and "location tracking" were appraised as attractive, motivating, and interesting. An "automatic activity recognition" feature was perceived as useful only under the condition that its precision was high. The "reminders" component was also deemed as useful only if a range of conditions was fulfilled (timeliness, opportunity for customization, etc). The features "mood and sleep tracking," "sharing workout results via social networks," "digital avatar and coach," and "rewards" were generally perceived negatively and considered as useless and not motivating. In general, participants preferred features with an easy-to-navigate interface and a clear, simplistic, and straightforward layout with a modern design. Customization and personalization qualities were highly appreciated throughout an app, together with data precision.

Conclusions: This study contributes to the understanding of the features and components preferred by adolescents in apps promoting PA. Such apps should provide users with precise data, and have a simplistic modern design and a straightforward easy-to-use interface. Apps should be personalized and customizable. Desired features to be included in an app are goal setting

and planning, feedback, coaching and training programs, and activity tracking. The features should involve high levels of data precision and timely delivery while taking into consideration the real-life context.

(*JMIR Hum Factors* 2022;9(2):e33972) doi:[10.2196/33972](https://doi.org/10.2196/33972)

KEYWORDS

mHealth; physical activity; mobile phone; health; qualitative research; focus groups; smartphone apps; behavior change; mobile health; adolescents

Introduction

Physical activity (PA) has been shown to be beneficial for both mental and physical health, while a lack thereof is known to be a risk factor for chronic and cardiovascular diseases [1]. Sufficient levels of PA in childhood and adolescence are particularly important, as it can set the standards for PA levels in adulthood [2]. The latest reports show that less than 20% of girls and 25% of boys reach the recommended levels of PA in European Union countries at the age of 15 years [1,3]. In view of the scale of the problem, it is crucial to develop interventions that promote and support PA in adolescents globally. Adolescents aged 16-18 years deserve special attention, as they display the lowest absolute PA levels among adolescents aged 5-19 years and are considered an at-risk group [4,5].

Smartphones have become ubiquitous devices among the young population over the last few years. In 2019, an estimated 94% of European young people accessed the internet on a daily basis, and 92% used mobile phones to access the internet away from home and work, according to Eurostat [6]. Considering their low implementation costs and pervasive presence, smartphones could be advantageous as a part of PA interventions, while also benefiting from multiple built-in sensors (eg, accelerometer, pedometer, GPS sensor, camera, and microphone) [7].

Smartphone apps are available through popular digital distribution services or app markets (eg, Apple App Store, Google Play Store, and Windows Phone Store). A recent review of apps aiming at improving diet, increasing PA, and reducing sedentary behavior in children and adolescents suggested that (1) there are fewer apps developed for adolescents than for adults; (2) the quality of apps is moderate, scoring the lowest for information quality and demonstrating a lack of theory-based (behavior change techniques and theories) and evidence-based (PA guidelines) approaches, which is congruent with reviews of PA apps for adults; (3) more formative research is needed to better understand the factors that improve adolescents' engagement and app quality [8]. Despite the increased interest in this field, there is a paucity of studies exploring and developing smartphone apps that promote and support PA in adolescents. A recent scoping review analyzing a range of evidence (both quantitative and qualitative) available on smartphone-based mobile health (mHealth) PA interventions and looking into the development and evaluation trajectory of smartphone-based mHealth PA interventions identified a lack of qualitative and quantitative studies exploring adolescents' views and experiences of apps promoting PA [9].

Existing qualitative studies exploring smartphone use for health purposes mainly focused on adult populations in several

contexts, including PA [10-13], health behavior change [14,15], health and fitness [16], and well-being [17]. Sample sizes in these studies vary, and the methods of data collection range from focus groups to online surveys, interviews, and "think aloud" methods. For example, Rabin and Bock [10] used a combination of a survey and a semistructured interview to collect feedback on 3 PA apps that can guide the development of theory-based and empirically based apps incorporating preferences of adults. Ehlers and Huberty [11] used an online survey to identify theory-based behavioral and technological features preferred by middle-aged women. Middelweerd et al [12] used a series of focus group discussions to explore students' preferences regarding a PA app. In their study, 30 participants aged between 18 and 25 years used the Nexercise app for 3 weeks and subsequently participated in a focus group discussion. Finally, Baretta et al [13] implemented a combination of a "think-aloud" methodology and in-depth interview techniques to examine the features of apps, such as "Runtastic Running & Fitness Tracker," "Endomondo - Running & Walking," and "Runkeeper - GPS Track Run Walk," which are important for users' engagement during the first exposure and after 2 weeks of using 1 of the 3 commercial apps.

In order to address these gaps in the literature, 2 separate focus group discussions were conducted in 2 groups of adolescents (n=4 each) aged 16-18 years, to investigate their experiences, attitudes, and preferences toward various PA app features and technologies that could (1) make the app more attractive for them and consequently (2) increase their interest and engagement with the app. This formative study was conducted to better understand the factors that improve adolescent engagement and app quality, and to ultimately inform the development of a mobile app focused on promoting PA among adolescents. The following research question was addressed: what features and components are preferred by adolescents (aged 16-18 years) in apps promoting PA?

Methods

Recruitment

As we intended to target a sample of adolescents aged 16-18 years, local schools were chosen as a recruitment location. As a consequence of lockdown measures because of the COVID-19 pandemic, it was difficult to recruit participants directly from schools, so we eventually used social networks as a recruitment platform. Students, who were interested and owned a smartphone, were initially asked to complete a questionnaire assessing their level of PA (Physical Activity Questionnaire for adolescents [PAQ-A]) [18,19], demographics, and experience with PA apps. Participants and their parents were asked to sign an informed consent form before taking part in the discussion

session. An effort was made to include participants with both high and low levels of PA, as individual PA profiles may affect the preference of certain features of PA apps [12]. The PAQ-A questionnaire score ranges from 1 (low PA) to 5 (high PA). After completing the questionnaire, participants were divided into the following 2 focus groups: a group with participants who had a PAQ-A score below 3 (low level of weekly PA), and a group with participants who had a PAQ-A score of 3 or above (moderate to high level of weekly PA). To ensure representativeness of the focus groups, additional attention was paid to gender balance in both groups.

Design

The design was guided by recommendations on the appropriate conduct of focus group discussions provided by Breen [20], and Krueger and Casey [21]. Both focus groups were carried out in Luxembourg online, using the videoconferencing software Skype for Business 2016. The moderator had previous experience of conducting qualitative research using interviewing techniques. Discussions were conducted using a semistructured interview guide. The moderator anticipated 90 minutes for each discussion; however, sessions could be prolonged, if needed. Both focus group discussions were audio and video recorded and transcribed verbatim, and the data were pseudonymized. A small incentive (€20 [US \$21] voucher) was sent to participants after their participation in the discussion session.

Ethics Approval

The study was approved by the Ethics Review Panel of the University of Luxembourg (ERP 19-046A2 MAPA).

Participants

Eligible participants (n=10) were students between 16 and 18 years of age. Due to dropout related to technical difficulties, only 8 participants took part in the focus group discussions. Participants were required to own a smartphone with internet access and to have some experience with PA apps prior to the session to ensure a meaningful focus group discussion. The sessions were conducted in English; therefore, all participants were required to have a sufficient command of the English language. Eligible participants were required to be healthy and to have no contraindications for PA participation.

Procedure

The 2 focus groups were held on separate days. Prior to the focus group discussions, signed consent forms were sent to the moderator. The moderator tested participants' language comprehension before commencing with the focus group discussion, and to avoid language barriers, the moderator used plain English and rephrased questions when needed. Then, during the online focus group session, the moderator welcomed participants and proceeded with the general overview of the topic. Further, the moderator ensured that participants were aware of the purpose of the study and its procedures, stated the ground rules for the focus group discussion, and underlined that the ensuing discussion was audio and video recorded for research purposes, assuring confidentiality and anonymity of transcriptions [20].

The focus group discussion started with a sequence of questions previously described by Dennison ([Multimedia Appendix 1](#)) [14]. Participants were initially asked to describe how often they use their smartphones, for which purposes, and which apps were used the most. In the next step, participants answered questions about their personal experience of smartphone apps for PA. To prompt a further discussion, the moderator used trigger materials that were explained and distributed among the participants ([Multimedia Appendix 1](#)). These materials included graphic examples of the most popular components present in both commercial and research-grade apps aimed at promoting PA. The list of app components presented to participants is shown in [Textbox 1](#). The moderator offered a summary of key questions and sought confirmation from participants.

After the interviewer presented each app component, participants were asked to comment on their thoughts and feelings in terms of perceived usefulness and relevance. Next, participants were asked to write their own "ideal" rewards that would motivate them to be more physically active in the chat window. At the final stage of the discussion, participants were asked questions about subsequent app development ("Questions specific for the MAPA app development trial" section in [Multimedia Appendix 1](#)). The analysis of this section was not included in this study. Although the focus group discussions were using a semistructured interview approach, sessions were conducted as "guided conversations," enabling the discussion to flow into unexpected directions [15]. At the end of the session, participants received the incentive as a token of gratitude.

Textbox 1. App components presented to the participants.

App components

1. Goal setting and planning
2. Coaching and training programs
3. Activity tracking
4. Mood and sleep tracking
5. Feedback
6. Sharing workout results via social networks
7. Social support and comparison (in-app social profile and challenges)
8. Location tracking
9. Automatic activity recognition
10. Digital avatar and coach
11. Rewards (virtual)
12. Reminders

Data Collection, Management, and Analysis

Both focus group discussions were video and audio recorded, and transcribed verbatim. A pseudonym was created for every participant to ensure anonymity. The transcripts were analyzed using the focus group discussion analysis methodology described by Krueger and Casey [21], which consist of categorizing participants' answers and comments according to the questions and themes from the focus group schedule using a word processor and consequently writing a descriptive summary for answers to each question or theme. A student assistant was involved in the transcription, coding, and categorization process. The fragments extracted from the transcripts were split according to the respective focus group discussion. This was done to identify differences between the 2 groups defined by their PA level. The data were later combined for further analysis [12].

Results

Overview

Among 10 eligible participants, 8 took part in focus group discussions (Table 1). Among the 8 participants, 2 were aged 16 years, 2 were aged 17 years, and 4 were aged 18 years (mean age 17.25 years, SD 0.82 years). To ensure gender balance, an equal number of males (n=4) and females (n=4) were enrolled. All participants were living in Luxembourg and studying in local schools (upper secondary education). All participants stated that they had no health issues or any other limitations preventing them from participating in any type of PA. Only 4 participants reported moderate to high levels of weekly PA, as assessed by the PAQ-A questionnaire. Every participant owned a smartphone and had experience of using at least one fitness app, mostly Garmin TrainingPeaks, Strava, Nike Training Club, or Adidas Runtastic, tracking mainly such activities as running, cycling, (gym) workouts, and swimming. Participants provided comments on various app components, which were further summarized as key themes.

Table 1. Participants' characteristics.

Focus group number	Age (years)	Gender	Grade	PAQ-A ^a score	Performed sports	Physical activity apps used
1	16	Female	3	2.75	Cycling and running	Garmin TrainingPeaks
1	18	Female	1	2.75	Running	Garmin TrainingPeaks and Strava
1	18	Female	1	2.6	Cycling, running, and rugby	Garmin TrainingPeaks, Strava, and Nike Training Club
1	16	Male	3	2.8	Cycling, running, soccer, and gym workouts	Garmin TrainingPeaks and Strava
2	18	Male	1	4.35	Cycling, running, soccer, and basketball	Adidas Runtastic
2	17	Male	2	3.5	Cycling, running, soccer, and basketball	Garmin TrainingPeaks
2	18	Female	1	3.68	Cycling, running, swimming, and volleyball	Adidas Runtastic, Garmin TrainingPeaks, and Nike Training Club
2	17	Male	2	3	Cycling, running, soccer, and skateboarding	Garmin TrainingPeaks

^aPAQ-A: Physical Activity Questionnaire for adolescents.

Ubiquitous Themes

Throughout the majority of discussed topics in both focus groups, certain comments indirectly related to the initial questions resurfaced on multiple occasions. These comments were united into “ubiquitous themes” related to app features. These themes were centered around design, customization and personalization, and data precision.

Design

Participants preferred a clear and simplistic app layout with an easy-to-navigate interface. Overall, apps were appraised while having modern looks and appealing colors (Adidas Runtastic and Garmin TrainingPeaks). Participants also preferred information to be displayed in a logical and straightforward manner. Any irrelevant or overcomplicated data were disliked, together with a scattered or complex layout.

Customization and Personalization

For an app to be appealing, it had to combine such qualities as flexibility and diversity. While customization refers to adjustments done by users, personalization refers to adjustments done by an app or a platform. Specifically, the majority of the features were perceived as appealing when the content was personalized and a user was able to customize it toward her/his preferences (hiding or unhiding various features, changing colors, etc). A user should be able to turn on/off different features (eg, location tracking), and the content (eg, proposed workouts) should be diverse and customizable.

Data Precision

Various app features were only perceived as useful if data provided in the feedback tab were precise and accurate; otherwise, they were perceived as useless and ineffectual.

Differences Between Focus Groups

The analysis showed agreement concerning general themes between both groups. The only difference between group 1

(participants with low PA levels) and group 2 (participants with moderate and high PA levels) concerned the perception of the “social support and comparison” feature (in-app social profile and challenges). This feature involves exposing user’s activity and achievements within an app’s social ecosystem. While group 1 appreciated this feature and perceived it as motivating and fun, group 2 mainly disliked it (accounting for the mental pressure this feature puts on participants).

So, during the quarantine we had like a running challenge where everyone who has this app could participate and the one with the most kilometers won. So, I really like that. It kind of motivated me to run because I didn't want to be last. [Group 1 participant, female]

I don't like it. I went running with a lot of people who actually have this fitness app that can really track other people that you follow. And they always ask me if I wanted to do it but I didn't, I just don't like it because I think that it puts a lot of pressure on you to be better than them and I want to run because I do it for myself and not for others so... [Group 2 participant, female]

General Themes

Goal Setting and Planning

This feature was generally perceived as useful and motivating (especially for those who do not train with a real-life coach). In general, participants enjoyed having the possibility to plan activities and have a clear outline of the activity agenda. They not only favored the possibility of choosing the preselected activities, but also enjoyed including their own activities as planning options.

Coaching and Training Programs

In general, this feature was perceived positively and appraised as especially useful for beginners who do not have a real-life coach. Participants outlined the motivational value of a coach

(real-life or online, eg, a familiar athlete), as well as the importance of flexibility and variability in online training programs. They enjoyed a more personalized interface and the ability to customize their training plan. Furthermore, they enjoyed reviewing the workout time frame and the training sample videos. Concurrently, the participants appraised some of the information as not useful (eg, calories burned during certain workouts). They outlined the importance of indicating the difficulty level and the equipment required for workout sessions.

I also like that the app has different options... And I also think the calories stuff is useless because it changes from person to person. So, it could be better if they told you whether it's a hard workout or just an easy one. [Group 2 participant, male]

Activity Tracking

Participants appreciated the activity tracking feature, specifically for such components as location, pace, duration, and distance tracking. Moreover, they proposed to integrate the audio player with the tracking interface (in order to simultaneously track the activity and listen to music) and underlined that an app has to request users for permission to track their location.

I like all of them because they all track the location and duration of the run, the distance. And there's nothing too, too complicated about it. You can just go with the map or listen to music also. I think it's yeah, it's a good thing to do. [Group 1 participant, female]

Mood and Sleep Tracking

Overall, participants did not perceive this feature as important and consequently gave mixed feedback. In general, they liked the idea of having data about sleep duration and mood assessment, yet only if it was precise. Participants shared their past (mostly negative) experiences and underlined that when such information is inaccurate, it is useless and they would not benefit from it.

I used lots of things, in the beginning when I had my smart watch for the first time. And actually, noticed that it isn't accurate. So, it said, I go to sleep at 9 o'clock when I actually go to sleep at 11 clock, so... [Group 2 participant, female]

Feedback

This feature was perceived very positively, but only on the condition of appropriate design (clear, simplistic, customizable, and only relevant information). Participants enjoyed the opportunity to review their monthly, weekly, or daily achievements.

I think that's probably one of the most important parts of an app. Because you can see all your data. Yeah so it needs to be very clean. Yeah. Not too much going on. [Group 2 participant, male]

I like the color (Garmin app). I like that you're running once a day and activity is in orange. So, it jumps out. So, it's the first thing you see. You have

the most important information next to it, so... I don't really like the other one [Fitbit app]. It's just too much. I don't need to see how many floors I climbed that day. It's just a bit useless. [Group 1 participant, female]

Sharing Workout Results via Social Networks

The majority of participants disliked this feature and would not want to share their activity results via social networks. They suggested this feature to be an optional component for other people who are active on social media but not for everyone.

Location Tracking

This feature was appraised positively and was useful for tracking various PAs. It was used for navigation purposes or for discovering and exploring new areas near home and during vacation. However, its function to run in the background throughout the day was not appreciated. Participants outlined that there has to be a clear way to turn the tracking feature off, and an app has to notify its users about tracking their location.

Automatic Activity Recognition

In general, participants found this feature useful only if the precision level was high; otherwise, it was considered unnecessary and useless according to participants past experiences.

I don't think it's very like precise or anything. But I don't really mind it. I think it can be useful if it's like precise. I mean I don't really need to know how much or how many minutes I walked. But yeah I just think it might be like fun to like look what they recognize as an exercise... [Group 1 participant, female]

Digital Avatar and Coach

In essence, this feature was perceived negatively and assessed as not motivating, useless, and unprofessional (associated rather with a game than a PA app). It was regarded to be useful for other people that perform home workouts. It was proposed to have this feature as an optional component.

Rewards (Virtual)

This feature was generally disliked and described as childish, not motivating, and not evoking feelings of pride. The monetization did not make sense to participants, and they disliked the shift in focus from internal to external motivation. Participants recommended it as an optional feature for beginners. The ideal rewards suggested were items related to preferred activities or workouts (eg, equipment, clothes, in-app upgrades, and a possibility to unlock premium features).

Reminders

This feature received mixed feedback and was deemed as a useful feature only in particular cases (eg, for beginners or other people not training with a real-life coach and in combination with an already established training program). In addition, it was perceived as motivating only if it could be customized, could be turned off when needed, and would appear only when relevant (eg, not during lessons or after the training).

Discussion

Principal Findings

This exploratory study aimed at investigating the preferences and attitudes of adolescents (aged 16-18 years) toward various PA app features and technologies that could potentially make an app more attractive and consequently increase interest and engagement.

Features, such as “goal setting and planning,” “coaching and training programs,” “activity tracking,” “feedback,” and “location tracking,” were preferred by focus group participants, and were appraised as attractive, motivating, and interesting. The “automatic activity recognition” feature was perceived as useful only under the condition that its precision is high. The “reminders” component was also deemed as useful only if a range of conditions was fulfilled (timeliness, opportunity for customization, etc). The features “mood and sleep tracking,” “sharing workout results via social networks,” “digital avatar and coach,” and “rewards” were generally perceived negatively and were considered as useless and not motivating. In general, participants preferred when features had an easy-to-navigate interface and had a clear, simplistic, and straightforward layout with modern design. Customization and personalization qualities were highly appreciated throughout an app, together with data precision.

While the comparison of focus group participants with low PA levels and those with moderate to high PA levels showed agreement in the majority of app features, the groups differed in their preference for the “social support and comparison” component in that the former liked it better than the latter. This difference suggests that “social support and comparison” may not be a primary feature in PA apps for adolescents.

These findings support previous research conducted with adults, with some exceptions. While in line with the work of Rabin and

Bock [10] confirming user preferences toward PA app features, such as user friendly interface, background music integration, goal setting, and tracking progress toward PA goals, our findings differ from the results reported by Ehlers and Huberty [11], who found that the most preferred technological features concerned components that enhance playfulness, competition with peers, and interaction in the app. The current findings are in line with most of the results reported by Middelweerd et al [12] in that users preferred a simple layout, the ability to tailor an app’s interface according to their needs, the tracking of PA using GPS, coaching features, tailored goals, and feedback. In contrast to these results, however, our participants had mixed thoughts concerning such features as competition with friends and a reward feature in the PA app. In addition, our findings are generally in line with the findings of Baretta et al [13], where features, such as simplicity, self-regulation skills support, and context tailoring, were perceived as important for users’ continuous engagement. Finally, when compared with the very few studies involving adolescents, our findings are in line with those of Lubans et al [22] and Seah and Koh [23], where features, such as goal setting, feedback, and activity tracking, were perceived as motivating by users.

Implications for Future Interventions

As this study was conducted exclusively with a Luxembourgish sample, conclusions must be drawn with caution as the results cannot be generalized to other populations. In addition, and subsequent to this qualitative approach, future studies should use a quantitative design involving a sufficiently powered sample of adolescents. Nevertheless, even at this preliminary stage, the results point toward the importance of a number of features concerning PA apps for adolescents. A preliminary list of recommended app features that researchers and developers may want to take into consideration when developing an app promoting and supporting PA in adolescents is presented in [Textbox 2](#).

Textbox 2. Preliminary list of recommended app features.

App features

1. Design: An up-to-date easy-to-navigate interface with a clear, simplistic, and straightforward layout is required. Features, such as customization and personalization, are appreciated, yet without overcomplicating the app interface.
2. Data: Information provided to users should be timely and precise. Auxiliary features indirectly linked to physical activity (PA) (eg, sleep duration and quality) may only be included when supported by precise data; otherwise, they may be excluded.
3. Goal setting and planning: It may be recommended for inclusion. This feature should provide the possibility to plan activities (generic activity suggestions should be proposed, yet should also be customizable) and have a clear outline of the activity agenda.
4. Feedback (on previously performed PA): It may be recommended for inclusion while providing the possibility for review of monthly, weekly, and daily achievements.
5. Coaching and training programs: It may be recommended for inclusion while assuring customization and variability of online training programs. The proposed programs may be supported by training sample videos and information about the required equipment.
6. Activity tracking: It may be recommended for inclusion. Information should be provided about location, pace, duration, and distance of the exercise (when possible). Moreover, an audio player integrated with the tracking interface (in order to simultaneously track the activity and listen to music) may be advantageous. It is important, however, to underline that users should be able to turn off this feature. In addition, the app should notify the user about tracking details and ask for permission from the user for location tracking.
7. Location tracking (while not exercising): It may be recommended for inclusion while discovering and exploring new exercise areas near the home or during vacation, but not for tracking the location in the background throughout the day. Similar to activity tracking, the user should be able to turn off this feature, and the app should notify the user about tracking details and ask for permission from the user for location tracking.
8. Automatic activity recognition: It may only be included if the app can provide precise results.
9. Mood and sleep tracking: It may only be included if the app can provide precise data.
10. Sharing workout results via social networks: It may only be included as an optional component.
11. Social support and comparison (in-app social profile and challenges): It may only be included as an optional component.
12. Digital avatar and coach: It may only be included as an optional component.
13. Rewards: It may only be included as an optional component for beginners.
14. Reminders: It may only be included if this feature is timely, relevant to the current context, and can be customized and turned off when needed.

Strengths and Limitations

To the best of our knowledge, this is the first qualitative study exploring adolescents' views on PA app features. These results should be replicated in future studies using quantitative designs and systematically investigating potential gender effects. The current results suggest several features of smartphone-based PA interventions for teenagers that should be considered by developers and researchers.

There are several limitations in this study. First, we did not provide participants with a specific mobile platform (app), rather participants reviewed screenshots from different PA apps. It could be argued that participants' assessments of app features may have differed when interacting with a functional app rather than viewing noninteractive screenshots. This should be replicated in future studies assessing participants' views of specific apps during use. Second, the current sample consisted of individuals who at some point were members of a sports club (most often specialized in running), and this may have affected the results. Moreover, as the sample was quite small, future researchers should confirm the findings using more representative samples, achieving better data saturation, and including a sufficient number of boys and girls to systematically investigate any gender effects. The findings of this study, therefore, cannot be generalized beyond similar populations because of its qualitative explorative characteristic.

Taking into account these limitations, the findings of this study provide the first evidence of teenagers' views on the features of PA promotion apps. It is hoped that this may stimulate future studies on larger and more representative samples, thereby providing conclusive evidence for developing effective PA promotion apps.

Conclusions

This study contributes to the understanding of the features and components preferred by adolescents in apps promoting PA. Such apps should provide users with precise data, and have a simplistic modern design and a straightforward easy-to-use interface. Apps should be personalized and customizable, than is, have the ability to be tailored toward users' needs and wishes. Desired features to be included in an app are goal setting and planning, feedback, coaching and training programs, and activity tracking. The features should involve high levels of data precision and timely delivery while taking into consideration the real-life context. This study provides initial information for both researchers and app designers working on the development of effective smartphone-based PA promotion interventions. Future quantitative studies should explore which app features could potentially increase motivation and improve long-term engagement of app users.

Acknowledgments

AD was supported by an Industrial Fellowship from the Luxembourg National Research Fund (reference number 12674722) and by Actimage Luxembourg SA. The authors are grateful for the support of the graduate psychology student Tabea Franziska Schmidt-Alkadri in the coding and categorization process.

Conflicts of Interest

None declared.

Multimedia Appendix 1

Focus group schedule and trigger materials.

[[PDF File \(Adobe PDF File\), 2896 KB - humanfactors_v9i2e33972_app1.pdf](#)]

References

1. OECD, European Union. Health at a Glance: Europe 2020: State of Health in the EU Cycle. Paris: OECD Publishing; 2020. URL: https://www.oecd-ilibrary.org/social-issues-migration-health/health-at-a-glance-europe-2020_82129230-en
2. OECD, European Union. Health at a Glance: Europe 2016: State of Health in the EU Cycle. Paris: OECD Publishing; 2016. URL: https://www.oecd-ilibrary.org/social-issues-migration-health/health-at-a-glance-europe-2016_9789264265592-en
3. WHO guidelines on physical activity and sedentary behaviour. World Health Organization. 2020. URL: <https://www.who.int/publications/i/item/9789240015128> [accessed 2022-05-25]
4. National Institute for Health Innovation. A National Survey of Children and Young People's Physical Activity and Dietary Behaviours in New Zealand: 2008/09. Ministry of Health. 2010. URL: <https://www.health.govt.nz/system/files/documents/publications/cyp-physical-activity-dietary-behaviours-08-09-keyfindgs.pdf> [accessed 2022-05-25]
5. Bhatti SN, Watkin E, Butterfill J, Li J. Recognition of 16-18-year-old adolescents for guiding physical activity interventions: A cross-sectional study. *Int J Environ Res Public Health* 2020 Jul 11;17(14):5002 [FREE Full text] [doi: [10.3390/ijerph17145002](https://doi.org/10.3390/ijerph17145002)] [Medline: [32664602](https://pubmed.ncbi.nlm.nih.gov/32664602/)]
6. Being young in Europe today - digital world. Eurostat. URL: https://ec.europa.eu/eurostat/statistics-explained/index.php?title=Being_young_in_Europe_today_-_digital_world [accessed 2022-05-25]
7. Harari GM, Lane ND, Wang R, Crosier BS, Campbell AT, Gosling SD. Using smartphones to collect behavioral data in psychological science: Opportunities, practical considerations, and challenges. *Perspect Psychol Sci* 2016 Nov;11(6):838-854 [FREE Full text] [doi: [10.1177/1745691616650285](https://doi.org/10.1177/1745691616650285)] [Medline: [27899727](https://pubmed.ncbi.nlm.nih.gov/27899727/)]
8. Schoeppe S, Alley S, Rebar AL, Hayman M, Bray NA, Van Lippevelde W, et al. Apps to improve diet, physical activity and sedentary behaviour in children and adolescents: a review of quality, features and behaviour change techniques. *Int J Behav Nutr Phys Act* 2017 Jun 24;14(1):83 [FREE Full text] [doi: [10.1186/s12966-017-0538-3](https://doi.org/10.1186/s12966-017-0538-3)] [Medline: [28646889](https://pubmed.ncbi.nlm.nih.gov/28646889/)]
9. Domin A, Spruijt-Metz D, Theisen D, Ouzzahra Y, Vögele C. Smartphone-based interventions for physical activity promotion: Scoping review of the evidence over the last 10 years. *JMIR Mhealth Uhealth* 2021 Jul 21;9(7):e24308 [FREE Full text] [doi: [10.2196/24308](https://doi.org/10.2196/24308)] [Medline: [34287209](https://pubmed.ncbi.nlm.nih.gov/34287209/)]
10. Rabin C, Bock B. Desired features of smartphone applications promoting physical activity. *Telemed J E Health* 2011 Dec;17(10):801-803. [doi: [10.1089/tmj.2011.0055](https://doi.org/10.1089/tmj.2011.0055)] [Medline: [22010977](https://pubmed.ncbi.nlm.nih.gov/22010977/)]
11. Ehlers DK, Huberty JL. Middle-aged women's preferred theory-based features in mobile physical activity applications. *J Phys Act Health* 2014 Sep;11(7):1379-1385. [doi: [10.1123/jpah.2012-0435](https://doi.org/10.1123/jpah.2012-0435)] [Medline: [24368818](https://pubmed.ncbi.nlm.nih.gov/24368818/)]
12. Middelweerd A, van der Laan DM, van Stralen MM, Mollee JS, Stuij M, te Velde SJ, et al. What features do Dutch university students prefer in a smartphone application for promotion of physical activity? A qualitative approach. *Int J Behav Nutr Phys Act* 2015 Mar 01;12:31 [FREE Full text] [doi: [10.1186/s12966-015-0189-1](https://doi.org/10.1186/s12966-015-0189-1)] [Medline: [25889577](https://pubmed.ncbi.nlm.nih.gov/25889577/)]
13. Baretta D, Perski O, Steca P. Exploring users' experiences of the uptake and adoption of physical activity apps: Longitudinal qualitative study. *JMIR Mhealth Uhealth* 2019 Feb 08;7(2):e11636 [FREE Full text] [doi: [10.2196/11636](https://doi.org/10.2196/11636)] [Medline: [30735143](https://pubmed.ncbi.nlm.nih.gov/30735143/)]
14. Dennison L, Morrison L, Conway G, Yardley L. Opportunities and challenges for smartphone applications in supporting health behavior change: qualitative study. *J Med Internet Res* 2013 Apr 18;15(4):e86 [FREE Full text] [doi: [10.2196/jmir.2583](https://doi.org/10.2196/jmir.2583)] [Medline: [23598614](https://pubmed.ncbi.nlm.nih.gov/23598614/)]
15. Miyamoto SW, Henderson S, Young HM, Pande A, Han JJ. Tracking health data is not enough: A qualitative exploration of the role of healthcare partnerships and mHealth technology to promote physical activity and to sustain behavior change. *JMIR Mhealth Uhealth* 2016 Jan 20;4(1):e5 [FREE Full text] [doi: [10.2196/mhealth.4814](https://doi.org/10.2196/mhealth.4814)] [Medline: [26792225](https://pubmed.ncbi.nlm.nih.gov/26792225/)]
16. Gowin M, Cheney M, Gwin S, Franklin Wann T. Health and fitness app use in college students: A qualitative study. *American Journal of Health Education* 2015 Jul 06;46(4):223-230. [doi: [10.1080/19325037.2015.1044140](https://doi.org/10.1080/19325037.2015.1044140)]
17. Chang T, Kaasinen E, Kaipainen K. What influences users' decisions to take apps into use?: a framework for evaluating persuasive and engaging design in mobile Apps for well-being. In: MUM '12: Proceedings of the 11th International

- Conference on Mobile and Ubiquitous Multimedia. 2012 Presented at: 11th International Conference on Mobile and Ubiquitous Multimedia; December 4-6, 2012; Ulm, Germany p. 1-10. [doi: [10.1145/2406367.2406370](https://doi.org/10.1145/2406367.2406370)]
18. Kowalski K, Crocker P, Donen R. The Physical Activity Questionnaire for Older Children (PAQ-C) and Adolescents (PAQ-A) Manual. University of Saskatchewan. 2004. URL: https://www.prismsports.org/UserFiles/file/PAQ_manual_ScoringandPDF.pdf [accessed 2022-05-25]
 19. Aggio D, Fairclough S, Knowles Z, Graves L. Validity and reliability of a modified english version of the physical activity questionnaire for adolescents. Arch Public Health 2016 Jan 22;74(1):3 [FREE Full text] [doi: [10.1186/s13690-016-0115-2](https://doi.org/10.1186/s13690-016-0115-2)] [Medline: [26807217](https://pubmed.ncbi.nlm.nih.gov/26807217/)]
 20. Breen RL. A practical guide to focus-group research. Journal of Geography in Higher Education 2006 Nov;30(3):463-475. [doi: [10.1080/03098260600927575](https://doi.org/10.1080/03098260600927575)]
 21. Krueger RA, Casey MA. Focus Groups: A Practical Guide for Applied Research. Thousand Oaks, CA: SAGE Publications, Inc; 2015. URL: <https://us.sagepub.com/en-us/nam/focus-groups/book243860>
 22. Lubans DR, Smith JJ, Skinner G, Morgan PJ. Development and implementation of a smartphone application to promote physical activity and reduce screen-time in adolescent boys. Front Public Health 2014;2:42 [FREE Full text] [doi: [10.3389/fpubh.2014.00042](https://doi.org/10.3389/fpubh.2014.00042)] [Medline: [24904909](https://pubmed.ncbi.nlm.nih.gov/24904909/)]
 23. Seah MLC, Koh KT. The efficacy of using mobile applications in changing adolescent girls' physical activity behaviour during weekends. European Physical Education Review 2020 Jun 17;27(1):113-131. [doi: [10.1177/1356336x20930741](https://doi.org/10.1177/1356336x20930741)]

Abbreviations

PA: physical activity

PAQ-A: Physical Activity Questionnaire for adolescents

Edited by A Kushniruk; submitted 01.10.21; peer-reviewed by A Middelweerd, S Stockwell; comments to author 31.12.21; revised version received 21.02.22; accepted 19.04.22; published 09.06.22.

Please cite as:

Domin A, Ouzzahra Y, Vögele C

Features and Components Preferred by Adolescents in Smartphone Apps for the Promotion of Physical Activity: Focus Group Study JMIR Hum Factors 2022;9(2):e33972

URL: <https://humanfactors.jmir.org/2022/2/e33972>

doi: [10.2196/33972](https://doi.org/10.2196/33972)

PMID: [35679113](https://pubmed.ncbi.nlm.nih.gov/35679113/)

©Alex Domin, Yacine Ouzzahra, Claus Vögele. Originally published in JMIR Human Factors (<https://humanfactors.jmir.org>), 09.06.2022. This is an open-access article distributed under the terms of the Creative Commons Attribution License (<https://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Human Factors, is properly cited. The complete bibliographic information, a link to the original publication on <https://humanfactors.jmir.org>, as well as this copyright and license information must be included.

Original Paper

Use of Health Care Chatbots Among Young People in China During the Omicron Wave of COVID-19: Evaluation of the User Experience of and Satisfaction With the Technology

Yi Shan¹, Prof Dr; Meng Ji², PhD; Wenxiu Xie³, MPhil; Xiaomin Zhang⁴, PhD; Xiaobo Qian⁵, BEng; Rongying Li⁶, BM; Tianyong Hao⁵, Prof Dr

¹School of Foreign Studies, Nantong University, Nantong, China

²School of Languages and Cultures, University of Sydney, Sydney, Australia

³Department of Computer Science, City University of Hong Kong, Hong Kong, China

⁴Department of Linguistics, Macquarie University, Sydney, Australia

⁵School of Computer Science, South China Normal University, Guangzhou, China

⁶School of Artificial Intelligence, South China Normal University, Guangzhou, China

Corresponding Author:

Yi Shan, Prof Dr
School of Foreign Studies
Nantong University
No. 9, Seyuan Rd
Nantong, 226019
China
Phone: 86 15558121896
Email: victorsyh@hotmai.com

Abstract

Background: Long before the outbreak of COVID-19, chatbots had been playing an increasingly crucial role and gaining growing popularity in health care. In the current omicron waves of this pandemic when the most resilient health care systems at the time are increasingly being overburdened, these conversational agents (CA) are being resorted to as preferred alternatives for health care information. For many people, especially adolescents and the middle-aged, mobile phones are the most favored source of information. As a result of this, it is more important than ever to investigate the user experience of and satisfaction with chatbots on mobile phones.

Objective: The objective of this study was twofold: (1) Informed by Deneche and Warren's evaluation framework, Zhu et al's measures of variables, and the theory of consumption values (TCV), we designed a new assessment model for evaluating the user experience of and satisfaction with chatbots on mobile phones, and (2) we aimed to validate the newly developed model and use it to gain an understanding of the user experience of and satisfaction with popular health care chatbots that are available for use by young people aged 17-35 years in southeast China in self-diagnosis and for acquiring information about COVID-19 and virus variants that are currently spreading.

Methods: First, to assess user experience and satisfaction, we established an assessment model based on relevant literature and TCV. Second, the chatbots were prescreened and selected for investigation. Subsequently, 413 informants were recruited from Nantong University, China. This was followed by a questionnaire survey soliciting the participants' experience of and satisfaction with the selected health care chatbots via wenjuanxing, an online questionnaire survey platform. Finally, quantitative and qualitative analyses were conducted to find the informants' perception.

Results: The data collected were highly reliable (Cronbach $\alpha=0.986$) and valid: communalities=0.632-0.823, Kaiser-Meyer-Olkin (KMO)=0.980, and percentage of cumulative variance (rotated)=75.257% ($P<0.001$). The findings of this study suggest a considerable positive impact of functional, epistemic, emotional, social, and conditional values on the participants' overall user experience and satisfaction and a positive correlation between these values and user experience and satisfaction (Pearson correlation $P<0.001$). The functional values (mean 1.762, SD 0.630) and epistemic values (mean 1.834, SD 0.654) of the selected chatbots were relatively more important contributors to the students' positive experience and overall satisfaction than the emotional values (mean 1.993, SD 0.683), conditional values (mean 1.995, SD 0.718), and social values (mean 1.998, SD 0.696). All the participants ($n=413$, 100%) had a positive experience and were thus satisfied with the selected health care chatbots. The 5 grade categories of participants

showed different degrees of user experience and satisfaction: Seniors (mean 1.853, SD 0.108) were the most receptive to health care chatbots for COVID-19 self-diagnosis and information, and second-year graduate candidates (mean 2.069, SD 0.133) were the least receptive; freshmen (mean 1.883, SD 0.114) and juniors (mean 1.925, SD 0.087) felt slightly more positive than sophomores (mean 1.989, SD 0.092) and first-year graduate candidates (mean 1.992, SD 0.116) when engaged in conversations with the chatbots. In addition, female informants (mean 1.931, SD 0.098) showed a relatively more receptive attitude toward the selected chatbots than male respondents (mean 1.999, SD 0.051).

Conclusions: This study investigated the use of health care chatbots among young people (aged 17-35 years) in China, focusing on their user experience and satisfaction examined through an assessment framework. The findings show that the 5 domains in the new assessment model all have a positive impact on the participants' user experience and satisfaction. In this paper, we examined the usability of health care chatbots as well as actual chatbots used for other purposes, enriching the literature on the subject. This study also provides practical implication for designers and developers as well as for governments of all countries, especially in the critical period of the omicron waves of COVID-19 and other future public health crises.

(*JMIR Hum Factors* 2022;9(2):e36831) doi:[10.2196/36831](https://doi.org/10.2196/36831)

KEYWORDS

health care chatbots; COVID-19; user experience; user satisfaction; theory of consumption values; chatbots; adolescent; youth; digital health; health care; omicron wave; omicron; health care system; conversational agent

Introduction

Background

Regretfully, more than 95% of the population suffers from particular health problems [1], and about 60% of them visit a doctor when merely affected by minor illnesses, including a cold, headache, and stomachache. Actually, 80% of these diseases can be cured with home remedies, without the intervention of a doctor [2]. In this scenario, health care chatbots are capable of monitoring people's health [1] by providing timely, useful health care information, especially during the omicron waves of COVID-19. These conversational agents (CA) play a crucial role in health care in the fast-paced world, where the public prefers to be addicted to social media rather than to be concerned about their health [3] and mobile phones are becoming the primary source of information. Meanwhile, chatbots are substantially alleviating the pressure on the already overloaded health care systems in various countries. Therefore, an upsurge in the development and application of health care chatbots has been witnessed since the advent of ELIZA in 1966, which served as a psychotherapist promoting communication with patients [4]. It inspired the design and application of other health care chatbots [5], including Casper [2], MedChat [2], PARRY [6], Watson Health [7], Endurance [7], OneRemission [8], Youper [9], Florence [10], Your.Md [11], AdaHealth [12], Sensely [13], and Buoy Health [14]. These leading chatbots offer patients tailored health and therapy information, recommended products and services, and personalized diagnoses and treatments based on confirmed symptoms [15]. Facing the repeated daunting waves of COVID-19, many people are craving information to respond to the coronavirus [16], which is incessantly mutating. This sudden surge in the demand for information is increasingly overtaxing health care resources [17], including various health care hotlines and clinic services, so health care chatbots seem to be the only possible solution [17,18]. Given the status quo, the user experience of and satisfaction with chatbots are more important now than ever before. Relevant studies have been undertaken in some countries to investigate the effectiveness [19], usability [20], and

acceptability [21]. Depending on technology acceptance theories (TAT), these studies on the use of health care chatbots focused on improving user experience and satisfaction through personalization [22], enjoyment [19], and novelty [23]. However, almost no investigation has been conducted in this respect among people in China from the perspective of the theory of consumption values (TCV).

Chatbots display unmatched advantages compared to other health care alternatives: alleviating the pressure on contact centers [24] and reducing contact-induced risks, satisfying unprecedented needs for health care information in the case of the shortage of qualified human agents [25], providing cost-effective 24/7 service [25], offering consistent service quality [26], and making no moral judgement on undesirable information provided by users [27]. The enhancement of these qualities motivates their increased use for health care purposes. This trend is being accelerated in the repeated outbreak waves of COVID-19, where chatbots are being used to screen potential infected cases [28], to help call centers to triage patients [29], and to recommend the most appropriate solutions to patients [29].

These selling points will facilitate popularizing health care chatbots only when the public is willing to utilize them and adopt their recommendations [30,31] in the face of the rampant COVID-19 pandemic. To promote adoption and adherence, many related studies have been undertaken in terms of the use of chatbots during this global health emergency to explore user reaction [32], to probe user experience and design considerations [33], to focus on the usage purposes [34], to identify differences in chatbot feature use by gender, race, and age [35], to improve the bot response accuracy [36], to investigate people's behavior when seeking COVID-19 information [37], and to introduce newly developed COVID-19-specific chatbots [38,39]. Apparently, few investigations [32,33] have examined the users' perception of these chatbots, but extant studies predominantly focus on technology acceptance [40,41], neglecting user experience and user satisfaction. Admittedly, user experience and user satisfaction are crucially significant because good user experience is the prerequisite of user adoption of information

systems (IS) [42,43] and user satisfaction is a crucial factor for IS acceptance intention [44,45].

To fight against the COVID-19 pandemic, chatbots have been used to provide psychological service for medical professionals and the general public in China [46]. Unfortunately, only 1 study, based on Deneche and Warren [47], investigated the user experience of and satisfaction with chatbots addressing COVID-19-related mental health in Wuhan and Chongqing, China [48]. However, this study focused on the determinants influencing user experience and satisfaction rather than on user experience and satisfaction per se. This gap in the literature needs to be filled.

Objective

The objective of this study was twofold: (1) Informed by Deneche and Warren's [47] evaluation framework, Zhu et al's [48] measures of variables, and the TCV [49], we designed a new assessment model for the user experience of and satisfaction with chatbots on mobile phones, and (2) we aimed to validate the newly developed model and use it to investigate the user experience of and satisfaction with the popular Chinese and English language chatbots for timely self-diagnosis and general information concerning COVID-19 and the latest virus variants among young people (aged 17-35 years) in China in order to provide evidence for the potential improvements and developments of chatbots to sustain adherence and adoption, which is undoubtedly an inevitable worldwide trend.

Based on the twofold research aim, we proposed the following hypotheses:

- Hypothesis 1: Explaining user behaviors in terms of diverse value-oriented factors (function, emotion, social influence, and environment), the newly developed comprehensive assessment model will have a high degree of reliability and validity and can better evaluate the user experience of and satisfaction with chatbots on mobile phones.
- Hypothesis 2: The informants will generally be satisfied with their experience of using popular health care chatbots.

Two facts justify the necessity of this research: Young people (aged 17-35 years), occupying a large portion of the population in China, are more addicted to mobile health care apps than other age groups, and sustainable user adoption of and adherence to chatbots in this population can considerably emancipate clinicians, enabling them to pay close attention to more complex tasks and enhance the availability of qualified health care services to the general public in China.

Methods

Overall Procedures

We followed 5 steps to reveal the user experience of and satisfaction with chatbots in young people (aged 17-35 years) in China. First, we established an assessment model evaluating user experience and satisfaction based on the related literature and TCV and designed a questionnaire according to the assessment model. Second, we screened and selected the chatbots to be investigated. Third, we recruited 413 students from Nantong University, China, as informants of this study.

Fourth, we collected the informants' demographic information, tested their health literacy, and solicited their experience of and satisfaction with the selected health care chatbots via a questionnaire survey. Finally, we conducted quantitative and qualitative analyses based on the data collected through the questionnaire.

Recruitment of Informants

Participants were recruited from among students of Nantong University, China. This university recruits around 8000 students annually, with the total number of students exceeding 30,000. On-campus psychological tests and students' counselors reported that a large percentage of students suffer from psychological problems of varying degrees during the repeat COVID-19 outbreaks. They urgently need intelligence-based CA for self-diagnosis and general information on the pandemic and the latest virus variants to ease their psychologically strained mind during the public health emergency. Their experience of and satisfaction with health care chatbots are, on the whole, representative and characteristic of the adolescent and middle-aged population in China. The questionnaire survey was approved and supported by the school authority in charge of students' affairs and the student participants themselves. It was conducted using the online questionnaire survey platform wenjuanxing [50] on January 8, 2022, and the survey lasted until no additional questionnaire was submitted online for 2 consecutive days (January 12, 2022). Over this period, the survey was announced to the entire student body of over 1000 at the School of Foreign Studies, Nantong University, through emails and WeChat groups. The reason informants were recruited from among these students is that only these English majors reach the English proficiency enabling them to experience the use of English language chatbots. Characteristic of all the schools of foreign studies of all colleges and universities in China, the overwhelming majority of students are female.

Selection of Health Care Chatbots

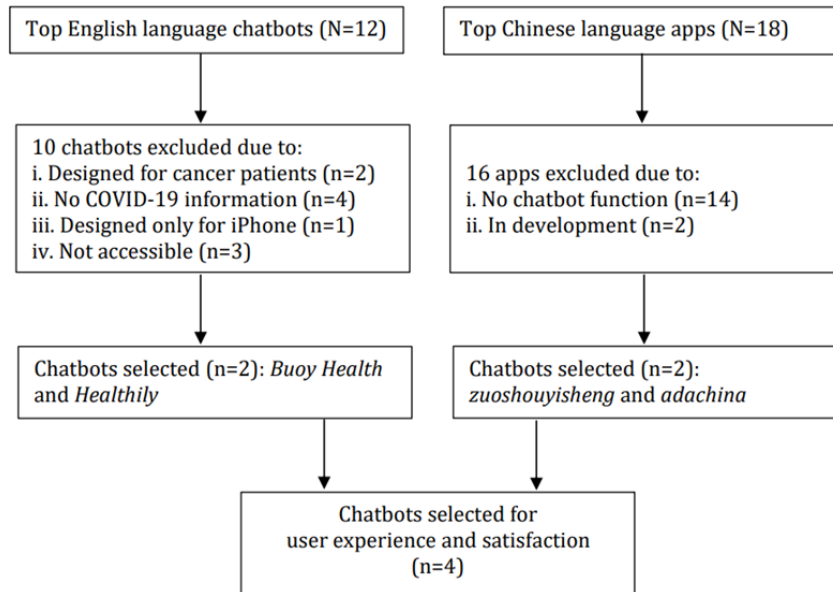
First, we chose the top 12 health chatbots popular throughout the world as the scope of selection of English language chatbots. These chatbots were reviewed by name, description, function, and experience, and only 2 (16.7%) of them, *Buoy Health* [14] and *Healthily* [11], were finally chosen (Figure 1).

Subsequently, we selected leading Chinese language chatbots from the dominant Android app markets, including *360 Mobile Assistant*, *Baidu Mobile Assistant*, and *Tencent MyApp*, and the iOS App Store. The keywords *health care chatbot* (医疗保健聊天机器人), *health care bot* (医疗保健机器人), *health care app* (医疗保健应用软件), *health care applet* (医疗保健小程序), *psychological health chatbot* (心理健康聊天机器人), *psychological health bot* (心理健康机器人), *psychological health app* (心理健康应用软件), and *psychological health applet* (心理健康小程序) were searched in Chinese on January 8, 2022. The selection followed 2 steps: (1) A total of 18 apps were identified by the search words, and (2) a further review revealed that only 4 (22.2%) of these 18 apps—*zuoshouyisheng* (左手医生), *adachina* (爱达健康), *zhinengyuwenzhen IPC* (智能预问诊IPC), and *xiaojiuzhinengwenzhenjiqiren* (小九智能问诊机器人)—have the chatbot function, while 2 (11.1%;

zhinengyuwenzhen IPC and *xiaojiuzhinengwenzhenjiqiren*) are still in development and provide no demos for experience and merely 2 (*zuoshouyisheng* and *adachina*) can truly function as chatbots. The selection process is illustrated in Figure 1.

Before answering the questionnaire, it was arranged that the informants would experience the use of both Chinese and English language chatbots for around 2 weeks. This 2-week experience was intended to guarantee the validity and reliability of the questionnaire survey.

Figure 1. Flowchart of selecting Chinese and English language health care chatbots. Of the top 12 English chatbots, 3 (25%) were not accessible due to technical errors, requirement of enterprise/school identification, or difficult application for a demo.



Assessment Model and Questionnaire

Informed by Deneche and Warren’s [47] evaluation framework, Zhu et al’s [48] measures of variables, and TCV [49], the assessment model designed for this study included 5 evaluation dimensions (functional, emotional, epistemic, social, and conditional) consisting of 18 variables (Table 1). These variables are supposed to contribute to user experience and user

satisfaction. The questionnaire included 36 measures (Multimedia Appendix 1). Measures 1-26 were designed in light of the variables listed in Table 1. To solicit sufficient information, some variables may have corresponded to more than 1 measure. For example, “performance” was related to 6 measures (15-20) in the questionnaire. Measures 27-36 were intended to display the informants’ overall experience and satisfaction.

Table 1. Assessment model of user experience and user satisfaction.

Dimension	Variables
Functional	<ul style="list-style-type: none"> Context awareness Language suitability Customized service User-friendliness Performance
Emotional	<ul style="list-style-type: none"> Enjoyment Relief from mental disorders
Epistemic	<ul style="list-style-type: none"> Novelty Desire for knowledge Knowledge enrichment
Social	<ul style="list-style-type: none"> Engagement Empathy Human likeness Privacy
Conditional	<ul style="list-style-type: none"> Time Place Technological context Mental state

Data Collection

The survey was conducted through wenjuanxing [50], an online questionnaire platform that is most popular in China. Three categories of data were collected via the online questionnaire: demographic information about the informants, their health literacy, and their experience of and satisfaction with the selected Chinese and English language chatbots. The demographic section collected data on the informants' age, gender, grade, English proficiency, and way to obtain health care information during the COVID-19 pandemic. The health literacy part tested the informants' basic medical vocabulary. The user experience and satisfaction module elicited data concerning the respondents' ratings of the 36 measures. The score of each measure was rated between 1 and 4 points (1: totally agree; 2: basically agree; 3: basically disagree; 4: totally disagree).

Data Analysis

Quantitative analyses were performed using SPSS Statistics version 22.0 (IBM Corp) and R version 4.0.2 (The R Foundation). First, the demographic data and health literacy of the participants were briefly described as the background information of the analysis. Afterward, the reliability and validity of the data concerning user experience and satisfaction were confirmed. Finally, the minimum, maximum, and mean scores, as well as SD, were calculated for each of the 36 measures, and the percentages of informants falling into each of the 4 ratings of the 36 measures were computed. Inspection of the data and residual plots for mean scores of the 36 measures did not indicate any violation of assumptions of normality, independence, and homogeneity of variance, so the correlation between measures 1-26 and measures 27-36 was tested and confirmed.

Ethical Considerations

Nantong University approved this study. It is an official practice in this university to ask the Students' Affairs Department for approval before collecting data from students. We followed this practice. In addition, there is no ethics review board in Nantong University. Therefore, a review number or code for this study could not be provided.

Results

Informant Demographics

A total of 413 questionnaires were collected, including 358 (86.68%) from female respondents. This can be explained by the fact that over 80% of students studying in the School of Foreign Studies, Nantong University, are female. The age of the participants ranged from 17 to 33 years (mean 20.96, SD 2.18). The overwhelming majority (n=402, 96.86%) of them are aged between 18 and 25 years. The informants included freshmen (n=66, 15.98%), sophomores (n=72, 17.43%), juniors (n=110, 26.63%), seniors (n=68, 16.46%), first-year graduate candidates (n=52, 12.59%), and second-year graduate candidates (n=45, 10.90%). They study in the School of Foreign Studies. Most of them (n=259, 62.71%) scored more than 100 in English in the entrance examinations for colleges and universities. Most of them (n=267, 65.65%) passed College English Test Band 6 (CET 6), Test for English Majors Band 4 (TEM 4), and TEM 8. Their English proficiency can well enable them to experience the use of English language chatbots. The majority of the informants (n=355, 85.96%) obtained COVID-19-related health care information through visiting a doctor or logging on to the internet. [Table 2](#) shows the informants' demographics, including grade, age, gender, and English proficiency, as well as the health care information sources they drew on.

Table 2. Informant demographics (N=413).

Categories	Participants, n (%)	Cumulative percentage (%)
I'm a ____.		
Freshman	66 (15.98)	15.98
Sophomore	72 (17.43)	33.41
Junior	110 (26.63)	60.05
Senior	68 (16.46)	76.51
First-year graduate candidates	52 (12.59)	89.10
Second-year graduate candidates	45 (10.90)	100.00
I'm ____ years old.		
17	2 (0.48)	0.48
18	27 (6.54)	7.02
19	56 (13.56)	20.58
20	89 (21.55)	42.13
21	79 (19.13)	61.26
22	65 (15.74)	77.00
23	47 (11.38)	88.38
24	19 (4.60)	92.98
25	18 (4.36)	97.34
26	3 (0.73)	98.06
27	2 (0.48)	98.55
29	1 (0.24)	98.79
32	3 (0.73)	99.52
33	2 (0.48)	100.00
I'm ____.		
Male	55 (13.32)	13.32
Female	358 (86.68)	100.00
I scored ____ in English in the entrance examinations for colleges and universities.		
>90	154 (37.29)	37.29
>100	79 (19.13)	56.42
>110	41 (9.93)	66.34
>120	75 (18.16)	84.50
>130	57 (13.80)	98.31
>140	7 (1.69)	100.00
I passed ____.		
CET ^a 3	52 (12.59)	12.59
CET 4	94 (22.76)	35.35
CET 6	47 (11.38)	46.73
TEM ^b 4	150 (36.32)	83.05
TEM 8	70 (16.95)	100.00
Facing COVID-19, I mainly obtain health care information through ____.		

Categories	Participants, n (%)	Cumulative percentage (%)
Visiting a doctor	94 (22.76)	22.76
Logging on to the internet	261 (63.20)	85.96
Reading books, papers, and journals	14 (3.39)	89.35
Families, friends, and classmates	31 (7.51)	96.85
Health care hotlines	4 (0.97)	97.82
Health care chatbots	9 (2.18)	100.00

^aCET: College English Test.

^bTEM: Test for English Majors.

Data Reliability and Validity

As shown in Table S1 in [Multimedia Appendix 2](#), Cronbach α (.986) for all the items (measures), except item (measure) 4, rated by all the 413 respondents was well above .9. If item 4 was deleted, Cronbach α increased merely by .001, so it was retained for the analysis. This indicates that the data collected for each measure in the questionnaire are highly reliable. The *corrected item-total correlation* for each measure was well above 0.4, which implies that the 36 measures are closely correlated.

The data were highly valid (Table S2 in [Multimedia Appendix 2](#)). The *communalities* for all the 36 items ranged from 0.632 to 0.823, well above 0.4, indicating that all these items are reasonable and should be included in the analysis. The value of Kaiser-Meyer-Olkin (KMO) value of 0.980 was substantially above 0.6, showing that all the data concerning the 36 items are suitable for extraction. The *percentage of variance (rotated)* for factors 1-3 was 30.428%, 28.077%, and 16.752%, respectively, and the *percentage of cumulative variance (rotated)* for the 3 factors was 75.257%, considerably above 50%. This means that all the data on all the items can be extracted validly.

User Experience and Satisfaction

[Table 3](#) displays the results of the descriptive analysis of user experience and satisfaction. The minimum, maximum, and mean scores were based on the rating scale of each measure (1: totally agree; 2: basically agree; 3: basically disagree; and 4: totally disagree). The mean scores of the 36 measures were lower than or slightly over 2, indicating that the respondents were inclined to totally or basically agree with these measures. In other words, they found the chatbots pleasurable and satisfactory in terms of the functional, emotional, epistemic, social, and conditional domains.

The functional domain displayed the lowest mean score (1.762, SD 0.630), closely followed by the epistemic domain (mean 1.834, SD 0.654). This indicates that the respondents were overall satisfied with the function of the selected chatbots when seeking self-diagnosis and general knowledge about COVID-19

and the latest virus variants and that they had enriched their COVID-19-related knowledge through the novel way of communication with the chatbots. The conditional, social, and emotional domains showed a similar mean score of slightly lower than 2. It follows that the participants found it necessary and technologically possible to obtain health care information through communicating with the chatbots via a mobile phone anytime and anyplace in the face of the rampant COVID-19 pandemic, which imposes on them mental stress in varying degrees. Additionally, they believed that seeking COVID-19-related health care information through communicating with the chosen chatbots was generally enjoyable and mentally relaxing and that the somehow humanlike empathetic chatbots made them socially and emotionally engaged in machine-human conversations. Furthermore, they basically thought that their personal information revealed in communication with the chatbots would be used for medical or research purposes rather than for unreasonable or even illegal ends. Overall, they had a pleasant and satisfactory experience when communicating with the chatbots for COVID-19-related self-diagnosis and health care information, as shown by the mean scores of *experience* (1.978, SD 0.639) and *satisfaction* (1.894, SD 0.617) in [Table 3](#).

[Table S3](#) in [Multimedia Appendix 2](#) shows the proportion of informants falling into each of the 4 ratings of the 36 measures. Over 80% (n=330) of the respondents totally and basically agreed with all measures, except measures 3, 4, 17, 23, and 28. Strikingly, more than 90% of the respondents totally and basically agreed with measures 5 (n=381, 92.25%), 7 (n=385, 93.22%), 11 (n=372, 90.07%), 14 (n=372, 90.08%), 15 (n=386, 93.46%), 18 (n=379, 91.77%), 31 (n=375, 90.80%), and 35 (n=388, 93.95%). Even for measures 3, 4, 17, 23, and 28, 312 (75.54%), 322 (77.97%), 320 (77.48%), 298 (72.15%), and 286 (69.25%) of participants totally and basically agreed, respectively. Specifically, the rates of students totally agreeing with the 36 measures ranged from 76 (18.40%) to 147 (35.59%) and those basically agreeing with these measures varied between 210 (50.85%) and 286 (69.25%). This means that most of the participating students showed a positive attitude toward their experience of the use of chatbots.

Table 3. Descriptive analysis of user experience and satisfaction. Items 1-36 represent the 36 measures in the questionnaire (N=413 for each item).

Item	Minimum score	Maximum score	Mean score (SD)	Median score
Conditional domain (mean 1.995, SD 0.718)				
1	1.000	4.000	1.908 (0.666)	2.000
2	1.000	4.000	1.971 (0.686)	2.000
3	1.000	4.000	2.048 (0.777)	2.000
4	1.000	4.000	2.051 (0.731)	2.000
Epistemic domain (mean 1.834, SD 0.654)				
5	1.000	4.000	1.738 (0.646)	2.000
6	1.000	4.000	2.000 (0.690)	2.000
7	1.000	4.000	1.765 (0.627)	2.000
Functional domain (mean 1.762, SD 0.630)				
8	1.000	4.000	1.978 (0.648)	2.000
9	1.000	4.000	1.891 (0.639)	2.000
10	1.000	4.000	1.881 (0.606)	2.000
11	1.000	4.000	1.881 (0.602)	2.000
12	1.000	4.000	1.942 (0.647)	2.000
13	1.000	4.000	1.927 (0.627)	2.000
14	1.000	4.000	1.862 (0.629)	2.000
15	1.000	4.000	1.794 (0.602)	2.000
16	1.000	4.000	1.932 (0.631)	2.000
17	1.000	4.000	2.046 (0.700)	2.000
18	1.000	4.000	1.872 (0.596)	2.000
19	1.000	4.000	1.896 (0.620)	2.000
20	1.000	4.000	1.891 (0.639)	2.000
Social domain (mean 1.998, SD 0.696)				
21	1.000	4.000	1.915 (0.639)	2.000
22	1.000	4.000	1.998 (0.695)	2.000
23	1.000	4.000	2.133 (0.775)	2.000
24	1.000	4.000	1.944 (0.675)	2.000
Emotional domain (mean 1.993, SD 0.683)				
25	1.000	4.000	1.976 (0.679)	2.000
26	1.000	4.000	2.010 (0.686)	2.000
Experience domain (mean 1.978, SD 0.639)				
27	1.000	4.000	1.913 (0.617)	2.000
28	1.000	4.000	2.155 (0.751)	2.000
29	1.000	4.000	2.019 (0.653)	2.000
30	1.000	4.000	1.913 (0.593)	2.000
31	1.000	4.000	1.891 (0.583)	2.000
Satisfaction domain (mean 1.894, SD 0.617)				
32	1.000	4.000	1.901 (0.593)	2.000
33	1.000	4.000	1.939 (0.634)	2.000
34	1.000	4.000	1.947 (0.648)	2.000
35	1.000	4.000	1.792 (0.595)	2.000

Item	Minimum score	Maximum score	Mean score (SD)	Median score
36	1.000	4.000	1.889 (0.617)	2.000

Correlation Between the 5 Domains and User Experience and Satisfaction

Table S4 in [Multimedia Appendix 2](#) also demonstrates that the 5 domains are intimately correlated with user experience and satisfaction, that is, the former considerably contributes to the latter. Pearson correlation was used to determine the correlation between each of the 26 measures (1-26) in the 5 domains and each of the 10 measures (27-36) in overall user experience and satisfaction. Statistics showed that each of the former 26 measures is positively correlated with each of the latter 10 measures, with $P < .001$ for each correlation and all correlation coefficients varying from 0.459 to 0.844. This indicates that the functional, epistemic, emotional, social, and conditional values of health care chatbots contribute positively to overall user experience and satisfaction, as far as the 413 informants of this study are concerned.

Differences in User Experience and Satisfaction by Gender and Grade

[Table 4](#) illustrates the mean scores of all the 36 measures rated by males and females. The *t* test revealed that there was a significant difference between male ratings and female ratings, with the former being significantly higher than the latter ($P < .001$), as shown in [Table 5](#). This implies that female participants were more positive in their experience of and satisfaction with health care chatbots compared to their male counterparts.

According to the *t* test ([Table 6](#)), there was a significant difference between freshmen's ratings and sophomores' ratings ($P < .001$), between freshmen's ratings and first-year graduate candidates' ratings ($P < .001$), between freshmen's ratings and

second-year graduate candidates' ratings ($P < .001$), between sophomores' ratings and juniors' ratings ($P = .004$), between sophomores' ratings and seniors' ratings ($P < .001$), between sophomores' ratings and second-year graduate candidates' ratings ($P < .001$), between juniors' ratings and seniors' ratings ($P < .001$), between juniors' ratings and first-year graduate candidates' ratings ($P = .01$), between juniors' ratings and second-year graduate candidates' ratings ($P < .001$), between seniors' ratings and first-year graduate candidates' ratings ($P < .001$), between seniors' ratings and second-year graduate candidates' ratings ($P < .001$), and between first- and second-year graduate candidates' ratings ($P = .002$). This indicates that freshmen had a better experience and greater satisfaction than sophomores, first-year graduate candidates, and second-year graduate candidates when communicating with health care chatbots for COVID-19-related information. Sophomores had a better experience and greater satisfaction than second-year graduate candidates but a less positive experience and lesser satisfaction than juniors and seniors. Juniors felt more positive than first- and second-year graduate candidates but less positive than seniors in their experience and satisfaction. Seniors had a better experience and greater satisfaction than first- and second-year graduate candidates. First-year graduate candidates felt more positive than second-year graduate candidates when engaged in conversations with the health care chatbots.

Overall, seniors were the most positive when expressing their experience of and satisfaction with health care chatbots, closely followed by freshmen and juniors. Slightly less positive, sophomores and first-year graduate candidates had similar experience and satisfaction. Second-year graduate candidates did not feel so positive as the other 5 grade categories.

Table 4. Mean scores of the 36 measures by gender.

Item	Mean score (SD) by males	Mean score (SD) by females
1	1.945 (0.803)	1.903 (0.643)
2	2.012 (0.782)	1.967 (0.671)
3	2.091 (0.867)	2.044 (0.764)
4	2.018 (0.805)	2.058 (0.720)
5	1.818 (0.772)	1.731 (0.624)
6	2.018 (0.828)	2.000 (0.667)
7	1.927 (0.813)	1.744 (0.591)
8	2.036 (0.769)	1.972 (0.628)
9	2.055 (0.826)	1.869 (0.602)
10	2.000 (0.816)	1.867 (0.566)
11	2.000 (0.754)	1.867 (0.576)
12	2.036 (0.838)	1.931 (0.613)
13	2.036 (0.838)	1.914 (0.587)
14	1.945 (0.780)	1.853 (0.603)
15	1.964 (0.860)	1.772 (0.549)
16	2.055 (0.848)	1.917 (0.590)
17	2.127 (0.862)	2.036 (0.673)
18	2.000 (0.793)	1.856 (0.558)
19	1.982 (0.871)	1.886 (0.573)
20	1.964 (0.816)	1.883 (0.608)
21	2.018 (0.871)	1.903 (0.595)
22	2.036 (0.881)	1.994 (0.663)
23	2.255 (0.927)	2.117 (0.749)
24	2.036 (0.793)	1.933 (0.655)
25	2.055 (0.826)	1.967 (0.654)
26	2.073 (0.836)	2.003 (0.661)
27	1.964 (0.769)	1.908 (0.591)
28	2.164 (0.918)	2.156 (0.723)
29	2.036 (0.793)	2.019 (0.630)
30	1.945 (0.756)	1.911 (0.565)
31	2.036 (0.769)	1.872 (0.547)
32	1.964 (0.816)	1.894 (0.552)
33	1.927 (0.790)	1.944 (0.607)
34	1.982 (0.805)	1.944 (0.621)
35	1.873 (0.795)	1.783 (0.559)
36	1.982 (0.805)	1.878 (0.583)

Table 5. Results of the *t* test of mean scores of the 36 measures by gender (*t* test $P < .001$).

Classification	Participants, n (%)	Minimum score	Maximum score	Mean score (SD)
Male	55 (13.32)	1.000	4.000	1.999 (0.051)
Female	358 (86.68)	1.000	4.000	1.931 (0.098)

Table 6. Results of the *t* test of mean scores of the 36 measures by grade.

Classification	Participants, n (%)	Minimum score	Maximum score	Mean score (SD)	Freshman <i>P</i> value	Sophomore <i>P</i> value	Junior <i>P</i> value	Senior <i>P</i> value	First-year graduate candidate <i>P</i> value	Second-year graduate candidate <i>P</i> value
Freshman	66	1.000	4.000	1.883 (0.114)	N/A ^a	<.001	.24	.08	<.001	<.001
Sophomore	72	1.000	4.000	1.989 (0.092)	<.001	N/A	.004	<.001	.81	<.001
Junior	110	1.000	4.000	1.925 (0.087)	.24	.004	N/A	.001	.01	<.001
Senior	68	1.000	4.000	1.853 (0.108)	.08	<.001	.001	N/A	<.001	<.001
First-year graduate candidate	52	1.000	4.000	1.992 (0.116)	<.001	.81	.001	<.001	N/A	.002
Second-year graduate candidate	45	1.000	4.000	2.069 (0.133)	<.001	<.001	<.001	<.001	.002	N/A

^aN/A: not applicable.

Discussion

Principal Findings

Young people aged 17-35 years constitute a population that is considered particularly receptive to health care chatbots during the omicron waves of COVID-19 for self-diagnosis and information about the latest virus variants. The findings of this study bring into focus the effect of the functional, epistemic, emotional, social, and conditional values of health care chatbots on the user experience and satisfaction of this specific population. Our findings suggest a considerable positive impact of these values on their overall user experience and satisfaction and a positive correlation between these values and user experience and satisfaction. By conducting an online questionnaire survey in the midst of the repeated outbreaks of the COVID-19 pandemic, we found that all the participants basically had a positive experience and were thus satisfied with the selected health care chatbots due to their generally satisfactory services. Results of the statistics also showed different degrees of experience of and satisfaction with the chosen health care chatbots among the 5 grade categories of participants: Seniors were the most receptive to health care chatbots for COVID-19 self-diagnoses and information, while second-year graduate candidates were the least receptive; freshmen and juniors felt slightly more positive than sophomores and first-year graduate candidates when engaged in conversations with the chatbots. In addition, female informants showed a relatively more receptive attitude toward the selected chatbots than male respondents. One possible reason for the relatively low reception among second-year graduate candidates is that they basically belonged to the oldest age group and were comparatively less willing to accept the novel way of obtaining information through communicating with chatbots. Although there are no studies devoted to age-related differences in user experience and satisfaction, this aspect deserves further investigation.

In addition to the chatbots' advantages, such as accessibility, cost-effectiveness, and flexibility [51], the functional, epistemic, emotional, social, and conditional values contributed to the overall pleasant experience and general satisfaction among the

413 respondents. According to statistics, the functional and epistemic values of the selected chatbots were the most important contributors to the students' positive experience and overall satisfaction. Functional values are concerned with functional and utilitarian performance [52]. In this study, the informants believed that the chatbots could be aware of the consulting context to use suitable language to provide personalized services based on their specific needs [53]. Personalization is a crucial function of artificial intelligence-based applications [54]. The selected chatbots of this study provided the survey participants with such personalized services as feedback, health reports, alerts, and recommendations [22], dealing with diverse mental health issues bothering different people during the repeated resurgences of COVID-19 [46] and leading to a higher level of user experience and satisfaction [22,55]. In addition, we found that other functional values, including user-friendliness, ease of use, and performance (eg, timely, precise, accurate, and effective answering, error-handling capacity) [47], also contributed to the participants' generally positive experience and overall satisfaction. Communicating with the health care chatbots offered student informants novelty and satisfied their desire for knowledge [49], too. The novel way of learning self-diagnoses and general information concerning COVID-19 and the latest virus variants led to a basically positive experience of and overall satisfaction with the health care chatbots on the part of the respondents. This is in tune with some extant studies [49,52,56].

The conditional, emotional, and social values played similar roles in providing the informants with good experience and general satisfaction. Facing numerous mental disorders caused by COVID-19 worldwide, people have suffered from stress, anxiety, depression, and other psychological problems [57]. As such, chatbots have been launched to psychologically assist people in many countries during COVID-19 [58]. Such particular conditions and situations of time, place, technology, and people's mental state [59,60] promote the decision [61] made by the informants to resort to health care chatbots for self-diagnosis and the general information about COVID-19 and the latest virus variants. The survey participants found that the health care chatbots were available almost anytime and

anyplace, providing faster health care services and reducing contact-induced risks. Thus, informed by Lee et al [62], we concluded that the conditional values of chatbots perceived by the participants in the face of the worldwide health emergency of COVID-19 positively influenced the user experience of and satisfaction with the health care chatbots. This finding is in line with recent studies [48,52].

As an emotional value of chatbots [48], enjoyment is an important element of chatbots [40]. The respondents of this study considered that communicating with the chatbots gave them an enjoyable feeling and considerably relieved them of stress, depression, and anxiety, as proven in recent studies [62,63]. The impact of enjoyment and delight on the user experience of chatbots [64], user adoption [65], and user satisfaction [19,66] has been proven by some studies. This feeling helped relieve the stress, depression, and anxiety [66] of the informants of this study during the critical period of repeated outbreaks of COVID-19, contributing to their positive experience of and overall satisfaction with the health care chatbots chosen for this research.

User experience during the human-product interaction results from all respects of user feelings (functional, emotional, social, etc) [67], each of which brings about a particular evaluation of the product or service concerned [68]. In this study, the social values of the health care chatbots were also perceived by the participants. They believed that the selected chatbots could fully engage them when they communicated with the chatbots for self-diagnosis and acquisition of general information concerning COVID-19 and the latest virus variants, thus satisfying their needs for communication, affection, and social belonging [69]. They thought that they felt the chatbots' empathetic tones when conversing about COVID-19-related health care information and that their personal information would not be misused unreasonably and illegally. Such humanlike empathy and privacy protection led to a more positive outlook, a feeling of emotional backup, and a sense of social belonging on the part of the informants, establishing trust and emotional connection between them and the chatbots [69].

Implications

Informed by Deneche and Warren's [47] evaluation framework, Zhu et al's [48] measures of variables, and TCV [49,70], this study established a new assessment framework to investigate the informants' user experience of and satisfaction with the selected health care chatbots. It advanced the theory regarding the user experience of and satisfaction with health chatbots from the perspective of TCV, enriching previous studies that focus little on this aspect [48]. Although previous studies have examined the user experience of and satisfaction with health chatbots in terms of effectiveness, usability, and acceptability, personalization, enjoyment, and novelty, they have explored this topic drawing on TAT [19-23,40,41,63], for example, the Technology Acceptance Model (TAM) and the Unified Theory of Acceptance and Use of Technology Model (UTAUT). TAM and UTAUT are primarily concerned with the relationship between the user behavior and the quality and function of technology-empowered products, so these theories fail to provide a full account of the utilization of health care chatbots

in various human-machine interaction settings, particularly in the context of the COVID-19-induced social distancing and even lockdown [48]. Comparing TAT with TCV, we found that the latter has a greater power of explanation: TCV comprehensively integrates a variety of value-oriented factors (functional emotional, epistemic, social, and conditional) into the account of the behaviors of users when engaging in communication with chatbots. Therefore, the user experience and satisfaction assessment model we established based on TCV is most likely to gain a better understanding of the user experience of and satisfaction with health care chatbots during the public health emergency of COVID-19 and other public health crises and natural disasters. In addition, the assessment scale of 36 items and 5 dimensions we newly developed is more comprehensive than Deneche and Warren's [47] international assessment framework and Zhu et al's [48] measures of variables, thereby having high reliability (Cronbach $\alpha=0.986$) and validity (KMO=0.980). Although many countries have provided chatbots to psychologically assist the public during the COVID-19-induced health emergency [58], almost no research has been conducted to study the user experience of and satisfaction with mental health chatbots during this pandemic [54]. This paper fills the gap in the extant literature.

On the practical facet, the new assessment framework of this research and the related findings can inspire artificial intelligence (AI) companies or scientific institutions to better design health care chatbots by giving top priority to the functional and epistemic values of these CAs while not neglecting their emotional, social, and conditional values. Health care chatbots integrating these 5 domains of values can enhance user experience and satisfaction. This paper also provides the governments of all countries with certain guidelines to choose and popularize health care chatbots in times of public health emergencies, such as COVID-19. As the first generation living with AI, we have the responsibility to design chatbots and make them ubiquitous and helpful to the whole society [69].

Limitations

Several limitations may influence the generalization of the findings reported in this paper. Most importantly, some of our findings may be biased due to the selection of respondents. The higher percentage of female respondents may be related to this bias. Particularly, the slightly higher level of user experience of and satisfaction with the selected health care chatbots may be attributed to the slightly higher percentage of female respondents. Additionally, we did not ask whether respondents had previous experience of health care chatbots, so we were unable to clarify whether our findings were biased by a mixture of respondents both with and without prior experience in this aspect. Finally, the survey is cross-sectional and lacks comparison to a period unaffected by the COVID-19 pandemic or to a different time of the year, and the data were collected merely from 1 university. We were unable to ascertain that the findings of this study can be generalized to the same age group in other regions or countries. The generalizability and validity of the findings and the assessment framework of this study need to be examined in further studies.

Conclusion

Government agencies worldwide have been providing the public with chatbots to psychologically assist them [58] in coping with a plethora of mental disorders caused by COVID-19 [57]. However, there is little focus on the user experience of and satisfaction with health care chatbots among young people in the literature. This study deals with the use of health care chatbots among young people (aged 17-35 years) in China, mainly investigating their user experience and satisfaction

through a newly designed assessment framework. The findings illustrate that the functional, epistemic, emotional, social, and conditional domains in the new assessment framework all have a positive impact on the participants' user experience and satisfaction. This paper advances the theory regarding the usability of health care chatbots, and chatbots for other purposes, enriching the literature. It also provides practical implications for chatbot designers and developers as well as for governments of all countries, especially in the critical period of the omicron waves of COVID-19 and other future public health crises.

Conflicts of Interest

None declared.

Multimedia Appendix 1

Survey questionnaire.

[DOCX File, 95 KB - [humanfactors_v9i2e36831_app1.docx](#)]

Multimedia Appendix 2

Supplementary tables.

[DOCX File, 71 KB - [humanfactors_v9i2e36831_app2.docx](#)]

References

1. Rahman M, Amin R, Liton M, Hossain N. Disha: an implementation of machine learning based Bangla healthcare chatbot. 2019 Presented at: 22nd International Conference of Computer and Information Technology; December 18-20, 2019; Dhaka, Bangladesh p. 18-20. [doi: [10.1109/iccit48885.2019.9038579](#)]
2. Bhirud N, Tataale S, Randive S, Nahar S. A literature review on chatbots in healthcare domain. *Int J Sci Technol Res* 2019;8(7):225-231.
3. Mathew R, Varghese S, Joy S, Alex S. 2019 Chatbot for disease prediction and treatment recommendation using machine learning. 2019 Presented at: 3rd International Conference on Trends in Electronics and Informatics; April 23-25, 2019; Tirunelveli, India p. 23-25. [doi: [10.1109/icoei.2019.8862707](#)]
4. Weizenbaum J. ELIZA—a computer program for the study of natural language communication between man and machine. *Commun ACM* 1966 Jan;9(1):36-45. [doi: [10.1145/365153.365168](#)]
5. Klopfenstein L, Delpriori S, Malatini S, Bogliolo A. The rise of bots: a survey of conversational interfaces, patterns, and paradigms. 2017 Presented at: Conference on Designing Interactive Systems; June 10-14, 2017; Edinburgh, UK p. 10-14. [doi: [10.1145/3064663.3064672](#)]
6. Colby KM, Weber S, Hilf FD. Artificial paranoia. *Artif Intell* 1971;2(1):1-25. [doi: [10.1016/0004-3702\(71\)90002-6](#)]
7. Adamopoulou E, Moussiades L. Chatbots: history, technology, and applications. *Mach Learn Appl* 2020 Dec;2:100006. [doi: [10.1016/j.mlwa.2020.100006](#)]
8. Oneremission: Making the Lives of Cancer Survivors Easier. URL: <https://keenethics.com/project-one-remission> [accessed 2022-01-06]
9. Youper. Expert Care for Anxiety and Depression. URL: <https://www.youper.ai/> [accessed 2022-01-06]
10. Florence: Your Health Assistant. URL: <https://florence.chat/> [accessed 2022-01-06]
11. Healthily: Your Trusted Guide to Health. URL: <https://www.livehealthily.com/> [accessed 2022-06-03]
12. Health. Powered by Ada. URL: <https://ada.com/> [accessed 2022-06-03]
13. Sensely: Increasing Access. Lowering Costs. Improving Health. URL: <https://www.sensely.com/> [accessed 2022-06-03]
14. When Something Feels Off, Buoy It. URL: <https://www.buoyhealth.com> [accessed 2022-06-03]
15. The Medical Futurist. The Top 12 Health Chatbots. URL: <https://medicalfuturist.com/top-12-health-chatbots> [accessed 2022-06-03]
16. Drees J. Led by COVID-19 Surge, Virtual Visits Will Surpass 1B in 2020: Report. URL: <https://www.beckershospitalreview.com/telehealth/led-by-covid-19-surge-virtual-visits-will-surpass-1b-in-2020-report.html> [accessed 2022-06-04]
17. Judson T, Odisho A, Young J, Bigazzi O, Steuer D, Gonzales R, et al. Implementation of a digital chatbot to screen health system employees during the COVID-19 pandemic. *J Am Med Inform Assoc* 2020 Jul 01;27(9):1450-1455 [FREE Full text] [doi: [10.1093/jamia/ocaa130](#)] [Medline: [32531066](#)]
18. Miner AS, Laranjo L, Kocaballi AB. Chatbots in the fight against the COVID-19 pandemic. *NPJ Digit Med* 2020 May 04;3(1):65 [FREE Full text] [doi: [10.1038/s41746-020-0280-0](#)] [Medline: [32377576](#)]

19. Fitzpatrick KK, Darcy A, Vierhile M. Delivering cognitive behavior therapy to young adults with symptoms of depression and anxiety using a fully automated conversational agent (Woebot): a randomized controlled trial. *JMIR Ment Health* 2017 Jun 06;4(2):e19 [FREE Full text] [doi: [10.2196/mental.7785](https://doi.org/10.2196/mental.7785)] [Medline: [28588005](https://pubmed.ncbi.nlm.nih.gov/28588005/)]
20. Cameron G, Cameron D, Megaw G, Bond R, Mulvenna M, O'Neill S, et al. Assessing the usability of a chatbot for mental health care. 2019 Presented at: International Workshop on Internet Science; April 2019; St. Petersburg, Russia. [doi: [10.1007/978-3-030-17705-8_11](https://doi.org/10.1007/978-3-030-17705-8_11)]
21. Ail MR, Rasazi Z, Mamun AA, Langevin R, Rawassizadeh R, Schubert L, et al. A virtual conversational agent for teens with autism: experimental results and design lessons. 2020 Presented at: Proceedings of the 20th ACM International Conference on Intelligent Virtual Agents (IVA'20); October 2020; Scotland, UK. [doi: [10.1145/3383652.3423900](https://doi.org/10.1145/3383652.3423900)]
22. Kocaballi AB, Berkovsky S, Quiroz JC, Laranjo L, Tong HL, Rezazadegan D, et al. The personalization of conversational agents in health care: systematic review. *J Med Internet Res* 2019 Nov 07;21(11):e15360 [FREE Full text] [doi: [10.2196/15360](https://doi.org/10.2196/15360)] [Medline: [31697237](https://pubmed.ncbi.nlm.nih.gov/31697237/)]
23. Nelson RR, Consoli D. An evolutionary theory of household consumption behavior. *J Evol Econ* 2010 Feb 6;20(5):665-687. [doi: [10.1007/s00191-010-0171-7](https://doi.org/10.1007/s00191-010-0171-7)]
24. Hao K. The pandemic is emptying call centers. AI chatbots are swooping in. *MIT Technology Review*. URL: <https://www.technologyreview.com/2020/05/14/1001716/ai-chatbots-take-call-center-jobs-during-coronavirus-pandemic/> [accessed 2022-06-04]
25. Mittal A, Agrawal A, Chouksey A, Shriwas R, Agrawal S. A comparative study of chatbots and humans. *Int J Adv Res Comput Commun Eng* 2016;5(3):1055.
26. AbuShawar B, Atwell E. Usefulness, localizability, humanness, and language-benefit: additional evaluation criteria for natural language dialogue systems. *Int J Speech Technol* 2016 Jan 4;19(2):373-383. [doi: [10.1007/s10772-015-9330-4](https://doi.org/10.1007/s10772-015-9330-4)]
27. Følstad A, Skjuve M. Chatbots for customer service: user experience and motivation. 2019 Presented at: Proceedings of the 1st International Conference on Conversational User Interfaces; 2019; New York, NY p. 1-9. [doi: [10.1145/3342775.3342784](https://doi.org/10.1145/3342775.3342784)]
28. Ross C. I asked eight chatbots whether I had Covid-19. The answers ranged from 'low' risk to 'start home isolation'. URL: <https://www.statnews.com/2020/03/23/coronavirus-i-asked-eight-chatbots-whether-i-had-covid-19/> [accessed 2022-06-04]
29. Ghosh S, Bhatia S, Bhatia A. Quro: facilitating user symptom check using a personalised chatbot-oriented dialogue system. *Stud Health Technol Inform* 2018;252:51-56.
30. Nadarzynski T, Miles O, Cowie A, Ridge D. Acceptability of artificial intelligence (AI)-led chatbot services in healthcare: a mixed-methods study. *Digit Health* 2019;5:2055207619871808 [FREE Full text] [doi: [10.1177/2055207619871808](https://doi.org/10.1177/2055207619871808)] [Medline: [31467682](https://pubmed.ncbi.nlm.nih.gov/31467682/)]
31. Gefen, Karahanna, Straub. Trust and TAM in online shopping: an integrated model. *MIS Quarterly* 2003;27(1):51. [doi: [10.2307/30036519](https://doi.org/10.2307/30036519)]
32. Dennis AR, Kim A, Rahimi M, Ayabakan S. User reactions to COVID-19 screening chatbots from reputable providers. *J Am Med Inform Assoc* 2020 Nov 01;27(11):1727-1731 [FREE Full text] [doi: [10.1093/jamia/ocaa167](https://doi.org/10.1093/jamia/ocaa167)] [Medline: [32984890](https://pubmed.ncbi.nlm.nih.gov/32984890/)]
33. You Y, Gui X. Self-diagnosis through AI-enabled chatbot-based symptom checkers: user experiences and design considerations. *AMIA Annu Symp Proc* 2020;2020:1354-1363 [FREE Full text] [Medline: [33936512](https://pubmed.ncbi.nlm.nih.gov/33936512/)]
34. Chu SY, Kang S, Yoo SC. The influences of perceived value of AI medical counseling chatbot service on the use intention: focused on the usage purpose of chatbot counseling of obstetrics and gynecology. *Health Serv Manag Rev* 2021;15(3):41-59. [doi: [10.18014/hsmr.2021.15.3.41](https://doi.org/10.18014/hsmr.2021.15.3.41)]
35. Schubel LC, Wesley DB, Booker E, Lock J, Ratwani RM. Population subgroup differences in the use of a COVID-19 chatbot. *NPJ Digit Med* 2021 Feb 19;4(1):30 [FREE Full text] [doi: [10.1038/s41746-021-00405-8](https://doi.org/10.1038/s41746-021-00405-8)] [Medline: [33608660](https://pubmed.ncbi.nlm.nih.gov/33608660/)]
36. Ghalebl M, Almutadha Y, Algarni F, Abdullah M, Felemban E, Alsharafi A, et al. Mining the chatbot brain to improve COVID-19 bot response accuracy. *Comput Mater Continua* 2022;70(2):2619. [doi: [10.32604/cmc.2022.020358](https://doi.org/10.32604/cmc.2022.020358)]
37. Skarpa PE, Garoufallou E. Information seeking behavior and COVID-19 pandemic: a snapshot of young, middle aged and senior individuals in Greece. *Int J Med Inform* 2021 Jun;150:104465. [doi: [10.1016/j.ijmedinf.2021.104465](https://doi.org/10.1016/j.ijmedinf.2021.104465)] [Medline: [33887589](https://pubmed.ncbi.nlm.nih.gov/33887589/)]
38. Sweidan S, Laban S, Alnaimat N, Darabkh K. SEG-COVID: a student electronic guide within Covid-19 pandemic. 2021 Presented at: 9th International Conference on Information and Education Technology (ICIET); March 2021; Okayama, Japan p. 27-29. [doi: [10.1109/iciet51873.2021.9419656](https://doi.org/10.1109/iciet51873.2021.9419656)]
39. Sweidan SZ, Abu Laban SS, Alnaimat NA, Darabkh KA. SIAAA - C: a student interactive assistant android application with chatbot during COVID - 19 pandemic. *Comput Appl Eng Educ* 2021 Apr 16;29(6):1718-1742. [doi: [10.1002/cae.22419](https://doi.org/10.1002/cae.22419)]
40. Ashfaq M, Yun J, Yu S, Loureiro SMC. I. Chatbot: modeling the determinants of users' satisfaction and continuance intention of AI-powered service agents. *Telemat Inform* 2020 Nov;54:101473. [doi: [10.1016/j.tele.2020.101473](https://doi.org/10.1016/j.tele.2020.101473)]
41. Luo X, Tong S, Fang Z, Qu Z. Frontiers: machines vs. humans: the impact of artificial intelligence chatbot disclosure on customer purchases. *Mark Sci* 2019 Sep 20;38(6):937-947. [doi: [10.1287/mksc.2019.1192](https://doi.org/10.1287/mksc.2019.1192)]
42. Deng L, Turner DE, Gehling R, Prince B. User experience, satisfaction, and continual usage intention of IT. *Eur J Inf Syst* 2017 Dec 19;19(1):60-75. [doi: [10.1057/ejis.2009.50](https://doi.org/10.1057/ejis.2009.50)]

43. Zhang B, Zhu Y. Comparing attitudes towards adoption of e-government between urban users and rural users: an empirical study in Chongqing municipality, China. *Behav Inf Technol* 2020 Mar 18;40(11):1154-1168. [doi: [10.1080/0144929x.2020.1743361](https://doi.org/10.1080/0144929x.2020.1743361)]
44. Li C, Fang Y. Predicting continuance intention toward mobile branded apps through satisfaction and attachment. *Telemat Inform* 2019 Oct;43:101248. [doi: [10.1016/j.tele.2019.101248](https://doi.org/10.1016/j.tele.2019.101248)]
45. Jung K, Lee D. Reciprocal effect of the factors influencing the satisfaction of is users. *APJIS* 1995;5(2):199-226.
46. Liu S, Yang L, Zhang C, Xiang Y, Liu Z, Hu S, et al. Online mental health services in China during the COVID-19 outbreak. *Lancet Psychiatry* 2020 Apr;7(4):e17-e18. [doi: [10.1016/S2215-0366\(20\)30077-8](https://doi.org/10.1016/S2215-0366(20)30077-8)] [Medline: [32085841](https://pubmed.ncbi.nlm.nih.gov/32085841/)]
47. Denecke K, Warren J. How to evaluate health applications with conversational user interface? In: Pape-Haugaard LB, Lovis C, Madsen IC, Weber P, Nielsen PH, Scott P, editors. *Digital Personalized Health and Medicine*. Amsterdam: European Federation for Medical Informatics (EFMI) and IOS Press; 2020:978-980.
48. Zhu Y, Janssen M, Wang R, Liu Y. is me, chatbot: working to address the COVID-19 outbreak-related mental health issues in China. User experience, satisfaction, and influencing factors. *Int J Hum-Comput Interact* 2021 Nov 01:1-13. [doi: [10.1080/10447318.2021.1988236](https://doi.org/10.1080/10447318.2021.1988236)]
49. Sheth JN, Newman BI, Gross BL. Why we buy and what we buy: a theory of consumption values. *J Bus Res* 1991 Mar;22(2):159-170. [doi: [10.1016/0148-2963\(91\)90050-8](https://doi.org/10.1016/0148-2963(91)90050-8)]
50. wenjuanxing. More Than Questionnaires/Online Exams. URL: <https://www.wjx.cn/> [accessed 2022-06-03]
51. Przegalinska A, Ciechanowski L, Stroz A, Gloor P, Mazurek G. In bot we trust: a new methodology of chatbot performance measures. *Bus Horiz* 2019 Nov;62(6):785-797 [FREE Full text] [doi: [10.1016/j.bushor.2019.08.005](https://doi.org/10.1016/j.bushor.2019.08.005)] [Medline: [22818403](https://pubmed.ncbi.nlm.nih.gov/22818403/)]
52. Teng C. Look to the future: enhancing online gamer loyalty from the perspective of the theory of consumption values. *Decis Support Syst* 2018 Oct;114:49-60. [doi: [10.1016/j.dss.2018.08.007](https://doi.org/10.1016/j.dss.2018.08.007)]
53. Xiao, Benbasat. E-commerce product recommendation agents: use, characteristics, and impact. *MIS Quarterly* 2007;31(1):137. [doi: [10.2307/25148784](https://doi.org/10.2307/25148784)]
54. Chen T, Guo W, Gao X, Liang Z. AI-based self-service technology in public service delivery: user experience and influencing factors. *Gov Inf Q* 2021 Oct;38(4):101520. [doi: [10.1016/j.giq.2020.101520](https://doi.org/10.1016/j.giq.2020.101520)]
55. Shi S, Wang Y, Chen X, Zhang Q. Conceptualization of omnichannel customer experience and its impact on shopping intention: a mixed-method approach. *Int J Inf Manag* 2020 Feb;50(4):325-336 [FREE Full text] [doi: [10.1016/j.ijinfomgt.2019.09.001](https://doi.org/10.1016/j.ijinfomgt.2019.09.001)] [Medline: [22818403](https://pubmed.ncbi.nlm.nih.gov/22818403/)]
56. El Qaoumi K, Le Masson P, Weil B, Ün A. Testing evolutionary theory of household consumption behavior in the case of novelty - a product characteristics approach. *J Evol Econ* 2017 Aug 26;28(2):437-460. [doi: [10.1007/s00191-017-0521-9](https://doi.org/10.1007/s00191-017-0521-9)]
57. Ransing R, Nagendrappa S, Patil A, Shoib S, Sarkar D. Potential role of artificial intelligence to address the COVID-19 outbreak-related mental health issues in India. *Psychiatry Res* 2020 Aug;290:113176. [doi: [10.1016/j.psychres.2020.113176](https://doi.org/10.1016/j.psychres.2020.113176)] [Medline: [32531628](https://pubmed.ncbi.nlm.nih.gov/32531628/)]
58. Smith AC, Thomas E, Snoswell CL, Haydon H, Mehrotra A, Clemensen J, et al. Telehealth for global emergencies: implications for coronavirus disease 2019 (COVID-19). *J Telemed Telecare* 2020 Mar 20;26(5):309-313. [doi: [10.1177/1357633x20916567](https://doi.org/10.1177/1357633x20916567)]
59. Omigie NO, Zo H, Rho JJ, Ciganek AP. Customer pre-adoption choice behavior for M-PESA mobile financial services. *IMDS* 2017 Jun 12;117(5):910-926. [doi: [10.1108/imds-06-2016-0228](https://doi.org/10.1108/imds-06-2016-0228)]
60. Pihlström M, Brush GJ. Comparing the perceived value of information and entertainment mobile services. *Psychol Mark* 2008 Aug;25(8):732-755. [doi: [10.1002/mar.20236](https://doi.org/10.1002/mar.20236)]
61. Hung C, Hsieh C. Searching the fit pattern between cultural dimensions and consumption values of mobile commerce in Taiwan. *Asia Pacific Manag Rev* 2010;15(2):147-165 [FREE Full text]
62. Lee S, Lee J, Kim H. A customer value theory approach to the engagement with a brand: the case of KakaoTalk Plus in Korea. *APJIS* 2018 Mar 30;28(1):36-60. [doi: [10.14329/apjis.2018.28.1.36](https://doi.org/10.14329/apjis.2018.28.1.36)]
63. Cheng Y, Jiang H. AI - powered mental health chatbots: examining users' motivations, active communicative action and engagement after mass - shooting disasters. *J Contingencies Crisis Manag* 2020 Sep 29;28(3):339-354. [doi: [10.1111/1468-5973.12319](https://doi.org/10.1111/1468-5973.12319)]
64. Rese A, Ganster L, Baier D. Chatbots in retailers' customer communication: how to measure their acceptance? *J Retail Consum Serv* 2020 Sep;56(3):102176. [doi: [10.1016/j.jretconser.2020.102176](https://doi.org/10.1016/j.jretconser.2020.102176)]
65. Kasilingam DL. Understanding the attitude and intention to use smartphone chatbots for shopping. *Technol Soc* 2020 Aug;62:101280. [doi: [10.1016/j.techsoc.2020.101280](https://doi.org/10.1016/j.techsoc.2020.101280)]
66. Abd-Alrazaq AA, Alajlani M, Alalwan AA, Bewick BM, Gardner P, Househ M. An overview of the features of chatbots in mental health: a scoping review. *Int J Med Inform* 2019 Dec;132:103978. [doi: [10.1016/j.ijmedinf.2019.103978](https://doi.org/10.1016/j.ijmedinf.2019.103978)] [Medline: [31622850](https://pubmed.ncbi.nlm.nih.gov/31622850/)]
67. Lallemand C, Gronier G, Koenig V. User experience: a concept without consensus? Exploring practitioners' perspectives through an international survey. *Comput Hum Behav* 2015 Feb;43:35-48. [doi: [10.1016/j.chb.2014.10.048](https://doi.org/10.1016/j.chb.2014.10.048)]
68. Yu M, Zhou R, Cai Z, Tan C, Wang H. Unravelling the relationship between response time and user experience in mobile applications. *Internet Res* 2020 May 15;30(5):1353-1382. [doi: [10.1108/intr-05-2019-0223](https://doi.org/10.1108/intr-05-2019-0223)]

69. Shum H, He X, Li D. From Eliza to XiaoIce: challenges and opportunities with social chatbots. *Front Inf Technol Electronic Eng* 2018 Jan 8;19(1):10-26. [doi: [10.1631/fitee.1700826](https://doi.org/10.1631/fitee.1700826)]
70. Sweeney JC, Soutar GN. Consumer perceived value: the development of a multiple item scale. *J Retail* 2001 Jun;77(2):203-220. [doi: [10.1016/s0022-4359\(01\)00041-0](https://doi.org/10.1016/s0022-4359(01)00041-0)]

Abbreviations

AI: artificial intelligence

CA: conversational agents

CET: College English Test

IS: information systems

KMO: Kaiser-Meyer-Olkin

TAM: Technology Acceptance Model

TAT: technology acceptance theories

TCV: theory of consumption values

TEM: Test for English Majors

UTAUT: Unified Theory of Acceptance and Use of Technology Model

Edited by A Kushniruk; submitted 27.01.22; peer-reviewed by D Xia, H Tanaka; comments to author 23.03.22; revised version received 25.03.22; accepted 14.05.22; published 09.06.22.

Please cite as:

Shan Y, Ji M, Xie W, Zhang X, Qian X, Li R, Hao T

Use of Health Care Chatbots Among Young People in China During the Omicron Wave of COVID-19: Evaluation of the User Experience of and Satisfaction With the Technology

JMIR Hum Factors 2022;9(2):e36831

URL: <https://humanfactors.jmir.org/2022/2/e36831>

doi: [10.2196/36831](https://doi.org/10.2196/36831)

PMID: [35576058](https://pubmed.ncbi.nlm.nih.gov/35576058/)

©Yi Shan, Meng Ji, Wenxiu Xie, Xiaomin Zhang, Xiaobo Qian, Rongying Li, Tianyong Hao. Originally published in *JMIR Human Factors* (<https://humanfactors.jmir.org>), 09.06.2022. This is an open-access article distributed under the terms of the Creative Commons Attribution License (<https://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in *JMIR Human Factors*, is properly cited. The complete bibliographic information, a link to the original publication on <https://humanfactors.jmir.org>, as well as this copyright and license information must be included.

Original Paper

Developing mHealth to the Context and Valuation of Injured Patients and Professionals in Hospital Trauma Care: Qualitative and Quantitative Formative Evaluations

Thymen Houwen^{1*}, MD; Miel A P Vugts^{2,3*}, PhD; Koen W W Lansink⁴, MD, PhD; Hilco P Theeuwes⁴, MD; Nicky Neequaye¹, MSc; M Susan H Beerekamp⁵, MD, PhD; Margot C W Joosen², PhD; Mariska A C de Jongh¹, PhD

¹Network Emergency Care Brabant, Elisabeth-TweeSteden Ziekenhuis, Tilburg, Netherlands

²Tranzo Scientific Centre for Care and Wellbeing, Tilburg School of Social and Behavioral Sciences, Tilburg University, Tilburg, Netherlands

³Public Health and Primary Care, Leiden University Medical Center, Leiden, Netherlands

⁴Department of Trauma Surgery, Elisabeth-TweeSteden Ziekenhuis, Tilburg, Netherlands

⁵Department of Trauma Surgery, Ziekenhuisgroep Twente, Almelo, Netherlands

*these authors contributed equally

Corresponding Author:

Thymen Houwen, MD

Network Emergency Care Brabant

Elisabeth-TweeSteden Ziekenhuis

Hilvarenbeekseweg 60

Tilburg, 5022 GC

Netherlands

Phone: 31 132212103

Email: t.houwen@etz.nl

Abstract

Background: Trauma care faces challenges to innovating their services, such as with mobile health (mHealth) app, to improve the quality of care and patients' health experience. Systematic needs inquiries and collaborations with professional and patient end users are highly recommended to develop and prepare future implementations of such innovations.

Objective: This study aimed to develop a trauma mHealth app for patient information and support in accordance with the Center for eHealth Research and Disease Management road map and describe experiences of unmet information and support needs among injured patients with trauma, barriers to and facilitators of the provision of information and support among trauma care professionals, and drivers of value of an mHealth app in patients with trauma and trauma care professionals.

Methods: Formative evaluations were conducted using quantitative and qualitative methods. Ten semistructured interviews with patients with trauma and a focus group with 4 trauma care professionals were conducted for contextual inquiry and value specification. User requirements and value drivers were applied in prototyping. Furthermore, a complementary quantitative discrete choice experiment (DCE) was conducted with 109 Dutch trauma surgeons, which enabled triangulation on value specification results. In the DCE, preferences were stated for hypothetical mHealth products with various attributes. Panel data from the DCE were analyzed using conditional and mixed logit models.

Results: Patients disclosed a need for more psychosocial support and easy access to more extensive information on their injury, its consequences, and future prospects. Health care professionals designated workload as an essential issue; a digital solution should not require additional time. The conditional logit model of DCE results suggested that access to patient app data through electronic medical record integration (odds ratio [OR] 3.3, 95% CI 2.55-4.34; $P < .001$) or a web viewer (OR 2.3, 95% CI 1.64-3.31; $P < .001$) was considered the most important for an mHealth solution by surgeons, followed by the inclusion of periodic self-measurements (OR 2, 95% CI 1.64-2.46; $P < .001$), the local adjustment of patient information (OR 1.8, 95% CI 1.42-2.33; $P < .001$), local hospital identification (OR 1.7, 95% CI 1.31-2.10; $P < .001$), complication detection (OR 1.5, 95% CI 1.21-1.84; $P < .001$), and the personalization of rehabilitation through artificial intelligence (OR 1.4, 95% CI 1.13-1.62; $P = .001$).

Conclusions: In the context of trauma care, end users have many requirements for an mHealth solution that addresses psychosocial functioning; dependable information; and, possibly, a prediction of how a patient's recovery trajectory is evolving. A structured development approach provided insights into value drivers and facilitated mHealth prototype enhancement. The findings imply

that iterative development should move on from simple and easily implementable mHealth solutions to those that are suitable for broader innovations of care pathways that most—but plausibly not yet all—end users in trauma care will value. This study could inspire the trauma care community.

(*JMIR Hum Factors* 2022;9(2):e35342) doi:[10.2196/35342](https://doi.org/10.2196/35342)

KEYWORDS

wounds and injuries; telemedicine; recovery of function; rehabilitation; patient care management; qualitative research; evaluation study; holistic health

Introduction

Background

Traumatic injuries impose a great physical, psychological, social, and economic burden on victims, relatives, and society. Globally, approximately 1 billion people need health care because of physical injuries [1]. Traumatic brain injuries (55.5 million) and spinal cord injuries (27 million) are the most prevalent types, together causing 17.6 million years of life lived with disability in 2016 [2]. Mostly and increasingly, injured people survive but are confronted with long-term rehabilitation and disabilities in the physical, emotional, and social domains [3,4]. Both in severe injuries and in less severe injuries, patients are at risk for developing symptoms of posttraumatic stress disorder (10%) or depression (7%) or become less productive in their work [3,5,6]. Returning to work is a driver in recovery trajectories, but it is often a lengthy or uncertain endeavor (ie, return to work success rates between 12% and 70% have been found) [7,8].

Health care providers face challenges to innovate their services, such as with mobile health (mHealth) app, to continuously reduce mortality rates as well as to strike the right balance between time and other resource investments and the optimization of patient experiences of health and service quality [9]. For example, *virtual fracture clinics* limit the use of resources in caring for patients with simple and stable fractures while ensuring consistent quality of care [9,10]. This model also appeared to be useful when the COVID-19 pandemic forced (orthopedic) outpatient clinics to limit face-to-face consultations [11]. Another concept aimed to prevent persistent pain symptoms after lower extremity injuries with web-based cognitive behavioral interventions supported by a nurse [12,13]. A preliminary randomized feasibility trial showed less pain intensity in comparison with the provision of an educational pamphlet. As an increasing number of people have internet access, websites, telemedicine, or mobile apps potentially assist in improving patient experiences, accessibility, and cost-effectiveness [14]. Despite the anecdotal success of eHealth and mHealth in trauma care, previous research also showed disappointing adoption, scale-up, spread, and sustenance of communication technologies in health care settings [15]. To prevent common pitfalls, the Center for eHealth Research and Disease Management (CeHRes) road map has been introduced as an evidence-based framework to develop eHealth solutions [16,17]. When technology attributes and features are more complex and stakeholder values are equivocal, the risk of failure increases [15]. Results of one research might not be generalizable across different populations with

non-self-selecting injured patients. In this view, failure to address facilitators and barriers in eHealth solutions was associated with unsuccessful implementation [18]. Therefore, facilitators and barriers should be mapped before developing new eHealth initiatives. By addressing eHealth development through an iterative and collective process of value propositions with all stakeholders, disappointing future impacts can be partially prevented. On the other hand, there is ample literature on the development of eHealth solutions wherein both professionals and patient users collaborate and are subjected to systematic needs inquiries, which is recommended to promote the uptake of eHealth innovations [16].

Therefore, inclusion of both groups could provide new insights in eHealth development.

Research Aims and Questions

In this study, we aimed to develop an mHealth app serving as a mode to deliver efficacious patient information and support that responds to important requirements of injured patients, health care professionals, and other stakeholders in a Dutch hospital trauma care setting. The CeHRes road map for development [16] was applied in anticipation of future implementation. Herein, we primarily focused on the perspectives of end users: injured patients and trauma care professionals. The objectives were to describe (1) experiences of unmet information and support needs of injured patients to promote their quality of life after hospitalization; (2) actual or expected barriers and facilitators according to trauma care professionals for the provision of information and support (with existing delivery modes or hypothetical innovative propositions) in their work context; and (3) drivers of value of web-based or mHealth apps for both key user groups, that is, the patients with trauma and trauma care professionals. By value, or utility, we mean perceptions (eg, usefulness, relative advantage, and expected outcomes), attitudes, or intentions antecedent to starting (eg, buying or adopting) or continuing app use [15,19,20].

Methods

Research and Development Design

Development steps were taken according to the CeHRes road map and included “contextual inquiry” (objective 1 and objective 2) and “value specification” (objective 3) [16]. Thus, user requirements and value drivers were established and prototyping was initiated. The scope of the reported steps is shown in [Figure 1](#). Key working principles were “stakeholder participation,” “entanglement of development and implementation,” and “continuous evaluation cycles” [16].

Stakeholders involved in the development process were patients with trauma as interview respondents or occasional team members, health care professionals which were mostly trauma surgeons and physiotherapists and in the same roles as patients, hospital information and communication technology (ICT) services, external software developers, the hospital privacy officer, the national association of trauma surgeons, project managers, and researchers. These stakeholders played a role

from inside or outside the *multidisciplinary team* (Figure 2). A core multidisciplinary team managed the processes of *steps of the road* in weekly meetings. User representatives occasionally participated when key decisions were made, either by attendance at team meetings with surgeons or separate consultation with patients. Other *external* stakeholders were involved in development in anticipation of future development steps or future implementation.

Figure 1. Research scope according to the Center for eHealth Research and Disease Management road map. Adapted from van Gemert-Peijnen et al [17]. The degree of transparency of the shapes (80%-20%) indicates the degree to which each iterative step was completed during the study period of the presented research. Ob: objective.

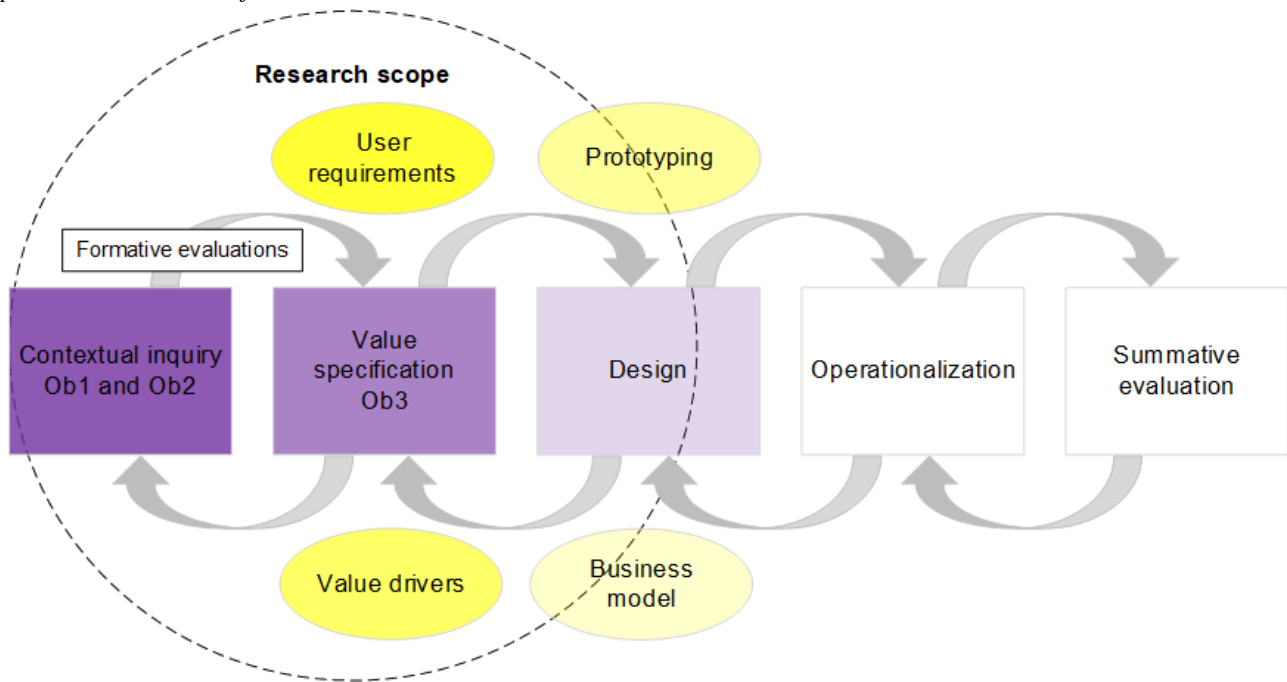
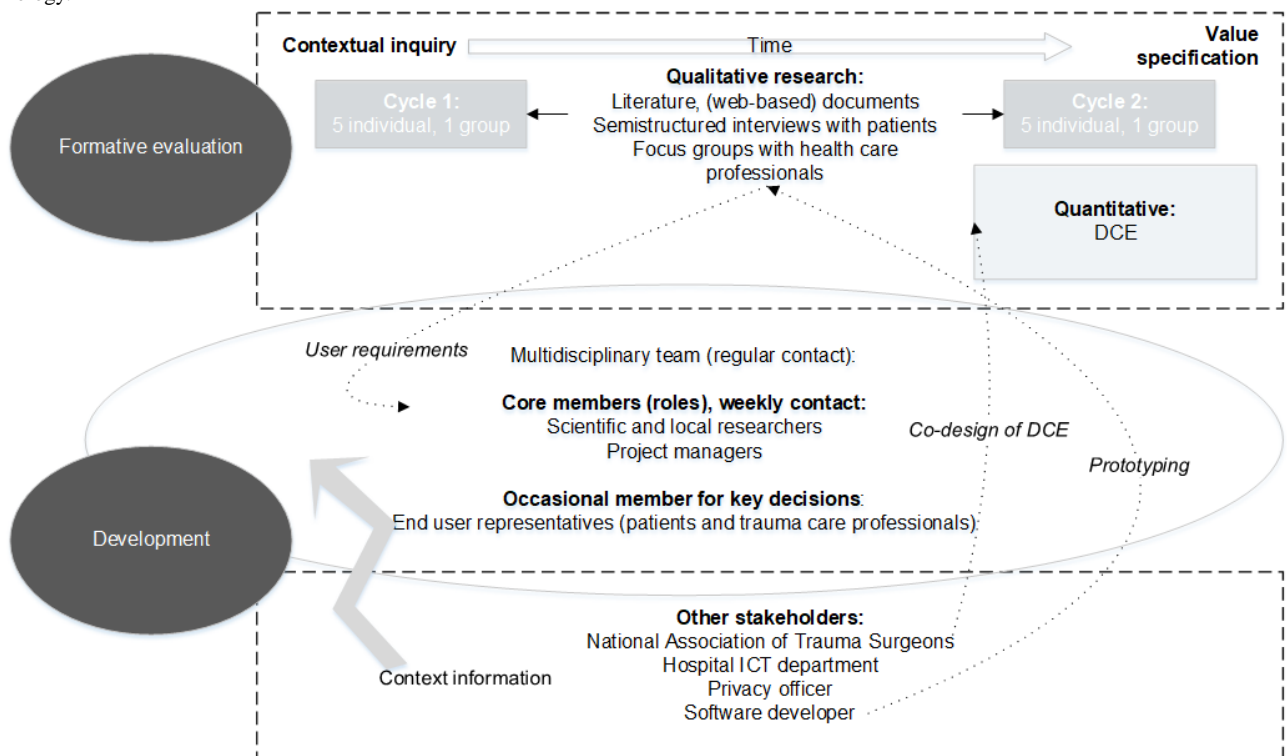


Figure 2. Structure of the development process and formative evaluation design. DCE: discrete choice experiment; ICT: information and communication technology.



A formative evaluation using qualitative and quantitative methods was conducted using a (partial) triangulation design [21]. The research methods included interviews with patients with trauma, a focus group with trauma care professionals, and a discrete choice experiment (DCE) among trauma surgeons.

First, qualitative research methods were used to accommodate the action-oriented, explorative, subjective, and in-depth nature of our objectives [22]. Individual patient interviews were completed in 2 cycles. The first interview cycle focused on user requirements and value drivers. These intermediate results were used to inform a phase of mobile app prototype development in collaboration with an external developer. In the second interview cycle, contextual fit and user value were explored using the developed mobile app prototype.

The focus group facilitated interactions between professionals outside routine situations to elicit complex thinking about implementing service changes involving patient information technologies.

Simultaneously, a complementary quantitative DCE with trauma surgeons was conducted to partly triangulate on results of the focus group on the topic of value specification (objective 3) [23]. Participating trauma surgeons stated their preference over hypothetical alternative mHealth products as described by a set of *attributes* and *levels*. For example, one attribute entailed “rehabilitation advice,” which could be *standardized* (level 0) or *personalized* with artificial intelligence (level 1). The DCE simulated trade-offs in deciding between 2 hypothetical mobile patient information apps with varying attributes and price levels. Value hierarchies (priorities) regarding the attributes of patient recovery technology could be inferred based on the chosen alternatives. In addition to an increase in sample size and representativeness (for a national) population of trauma surgeons, quantitative statistical modeling of user value attributions with a DCE provided a way of quantifying preferences and willingness to pay for attributes across decision makers. Ongoing development initiatives and previous qualitative insights obtained by members of an innovation committee from the Dutch Association of Trauma Surgeons (eg, SB) provided a unique window of opportunity for performing a DCE in this target group. There was no such opportunity to also conduct a DCE with patients.

Setting and Participants

The research took place in the Netherlands, where both surgical and orthopedic trauma care are part of the daily work of trauma surgeons. The qualitative research was conducted in a single level I trauma care center, Elisabeth-TweeSteden Hospital, the designated center for treating severely injured patients within its region. Eligible patients were of working age (18-67 years) and had a traumatic injury from 9 months to 5 years ago. The exclusion criteria were those with (1) a severe traumatic brain injury (ie, Glasgow Coma Score of <8), (2) dementia, or (3) insufficient command of the Dutch language. Focus groups were open to trauma surgeons and paramedics (physiotherapists) who provided direct care to patients from the Elisabeth-TweeSteden Hospital.

The DCE was conducted on a nationwide level, covering level I, II, and III trauma care centers. Dutch trauma surgeons were invited to participate in the DCE.

Qualitative Data Collection: Interviews With Patients With Trauma

A treating physician (TH, MD, male) selected eligible (former) patients from a trauma registration system and sent research invitations by email. Patient participants (via Microsoft Forms) completed screenings (age, gender, time since hospitalization, and work status) and provided informed consent—for a broader qualitative study. Then, accounting for respondents’ preferred ways of participating and aiming for maximal variation on the screening results, a purposive selection of candidates was sent interview invites, complementary informed consent documents, and screening questions on eHealth literacy and readiness. The participants were provided with a gift voucher of €40 (US \$43).

Patient interviews were semistructured and conducted by an experienced researcher (MAPV, PhD, male) via Microsoft Teams. Patients did not establish a relationship with the interviewer before starting the interviews. Each interview lasted 60 minutes and started with an introduction to this study. Next, open questions about experiences of traumatic events, received care and support, (unmet) information and support needs, and the suitability of various modes of delivery were asked. Field notes were taken during the interview. In the second part of each interview, patients were informed of several pre-existing and unevaluated ideas for eHealth or mHealth attributes to deepen their understanding of contextual fit and value considerations [24]. The topic list for the first cycle interviews included short explanations of potential attributes ([Multimedia Appendix 1](#)). During the second cycle of semistructured interviews, the list was replaced by a customized prototype containing all current ideas to meet patient user requirements. The prototype was shared on the screen by the interviewer, but patients could, if they wanted to, install and explore the prototype during days before the interview. Multiple screenshots of the prototype are provided in [Multimedia Appendix 1](#). No repeated interviews were conducted.

Qualitative Data Collection: Focus Groups With Trauma Care Professionals

We planned for 2 to 4 focus group of 90 minutes with 3 to 6 trauma care professionals in each group, which were facilitated by experienced researchers MCWJ (PhD, female) and MAPV via Microsoft Teams. Focus group participant selection targeted trauma health care professionals in the role of trauma surgeons or physiotherapists because of their systematic involvement in aftercare of patients with trauma. The planning of the group meetings adjusted to milestones of the development processes and circumstances related to COVID-19—measures that prevented meeting face-to-face, and time restraints and priorities of hospital staff limiting the opportunity to participate. After introducing the study background and aims, discussions focused on facilitators of and barriers to providing information and support and patient recipient and outcome specifications.

Qualitative Analysis: Interviews and Focus Groups

After the first interview cycle, one author (MAPV) immersed in the data, relistened the audio files, and made summaries to communicate the user requirements with the development team. After removal of personally identifiable information, verbatim audio transcriptions of all interview material were coded in pairs (TH and MAPV) using Atlas.ti 8 (ATLAS.ti Scientific Software Development GmbH). Relevant fragments of text content about the context and valuation of eHealth propositions were thematically coded using a combination of new labels (open codes) and pre-established codes (ie, sensitizing concepts) [25]. Before coding, a list of potential labels was established and piloted with the help of a third author (MCWJ). This list of potentially applicable codes included concepts from suitable frameworks to describe patients' health states as context (eg, labels about environmental factors and dimensions of functioning from the International Classification of Functioning and injury types from the International Classification of Diseases) and psychological constructs relating to the valuation of technology (perceptions, attitudes, or intentions toward technology use) [19,26]. These frameworks were chosen and discussed explicitly before coding for transparency and consistency. They reflected the complementary backgrounds of the coders in behavioral science (MAPV) and medicine (TH). Using these frameworks as part of the thematic analysis, we identified themes that were clearly defined and embedded in larger frameworks. Open codes were used when important data

could not be meaningfully labeled with the listed concepts. In the second coding step, codes were grouped with labels or higher theoretical abstraction, either building on existing frameworks or applying new inducted themes.
















Quantitative Data Collection and Analysis: DCE With Trauma Surgeons

First, the DCE was conducted according to recommended steps [27]. Additional details for each step are presented in Multimedia Appendix 2 [28]. Hypothetical app attributes and levels were established through a collaboration between researchers and the eHealth working group of the Dutch Association for Trauma Surgeons, represented by MSHB (MD and PhD). Six attributes were established with 2 to 4 different levels (Figure 3), which were refined based on feedback from the eHealth work group about a mock-up of the DCE.

Second, the constructed tasks involved taking repeated and free decisions between 1 out of 2 scenarios of hypothetical mobile patient information and support apps, or to *opt-out* (third option). Herein, opting out may have the relevant meaning of not adopting a newly (jointly) developed solution (*business-as-usual*) [29].

Third, the experimental design was specified in the R language (AlgDesign package) [30]. A total of 11 choice sets were made to adjust to respondents' limited time to concentrate on the decision tasks; issues with the data of response variability or censoring were considered plausible with a longer test [31].

Figure 3. Attributes and levels as in the instrument design. PROM: patient-reported outcome measure.

Attributes for trauma patients			
(Reference) level 1	Level 2	Level 3	Level 4
Generic content from the Dutch Association for Trauma Surgeons:  	Contact information and logo of own hospital visible. Generic information:  	Logo and local adjustment of information by local hospital  	
No complication detection system: 	Early complication detection, for example send wound pictures to the doctor 		
Standard rehabilitation according to a predetermined schedule 	Personalized rehabilitation via (self-learning) algorithms 		
No PROMS: 	Periodic self-measurements: PROMS, range of motion, or other 		
Attributes for trauma care professionals			
(Reference) level 1	Level 2	Level 3	Level 4
Patient data is accessible through a "standalone" website 	Patient data is viewed in a web viewer in Electronic medical record (EMR) system 	Patient data from the application is integrated in the EMR system 	
Yearly price: Free	Yearly price: < €2500 (< US \$2700)	Yearly price: €2500 (US \$2700) to €7500 (US \$8000)	Price: €7500 (US \$8000) to €15000 (US \$16000)

Fourth, preferences were elicited by informing participants about actual nationwide application development plans and explaining the DCE rationale. Before presenting the choice tasks, attributes and levels were explained as shown in [Figure 3](#). Choice tasks were presented by the question “which application has your preference?” and the answering possibilities “app 1,” “app 2,” or “none of the above apps.”

Fifth, MAPV and MSHB designed the instrument. Hypothetical apps were visually presented as a mobile phone screen with icons of attribute levels shown on 6 planes ([Figure 3](#)).

Sixth, the instrument was included in a survey on the priorities of trauma surgeons for patient information resources sent by the Dutch Association of Trauma Surgeons to all members. Similar to a previous DCE with health care professionals under similar circumstances, the board of the association consented to publishing DCE data without obtaining informed consent from individual participants [28]. Complementarily to the choice sets, survey information was used regarding participant characteristics.

Seventh, the panel data from the DCE were analyzed with conditional and mixed logit models using the *gaml* package in R [32]. Given the use of a convenience sample and a restricted number of choice tasks, sensitivity of the results was checked to various modeling approaches and settings. We have presented figures with results from both the conditional and mixed logit models as either choice has its own advantages and disadvantages. The conditional logit model is a good choice as it requires less data and the results are relatively easy to interpret. Mixed logit models often provide better results as they are more flexible with regard to modeling differences among decision makers but are more data demanding. To illustrate the implications of the modeling results for the valuation of mHealth apps with varying compositions, we calculated the predicted probabilities of selection for 4 different scenarios of attribute (level) combinations as the product of the relevant odds scores (*total odds*) divided by the sum of *total odds* and 1. Other details are presented in [Multimedia Appendix 2](#).

Ethics Approval

The Tilburg University ethical review board (RP301) approved this qualitative research.

Results

Recruitment

Patients

A total of 10 individual patient interviews were conducted with 5 male and 5 female patients, with a median age of 58 (range 21-67) years. An invitation was sent to 51 patients. Of the 10 patients, 5 (50%) were interviewed during the first interview cycle and the other 5 (50%) during the second cycle that included a demonstration of the mHealth app. The involved patients had various injuries such as a mild traumatic brain injury or complex or less complex bone fractures of the wrist

or ankle. The injuries were because of different traumatic causes (eg, road accident or activity-related injury). All but one were employed, and all were working for the same employer since their injury. The time since the injury ranged from <1 to 3 years. Furthermore, 90% (9/10) of participants regularly used a PC or laptop, 50% (5/10) used a tablet or iPad, and only 10% (1/10) did not use a mobile phone for internet browsing; 80% (8/10) of participants used mobile phones for SMS text messages. Except for 1 patient, all patients previously searched for additional health information and would use the internet for health information. All participants stated that they were sufficiently skillful to find helpful health resources on the internet.

For the second interview cycle, 5 patients evaluated the prototype of a mobile patient app.

Trauma Care Professionals

One focus group was conducted with 4 trauma care professionals of which 2 were male trauma surgeons and 2 were physiotherapists (1 male). All participants were working at a level I trauma center.

The respondents to the DCE consisted of 109 surgeons who provided entries for all presented choice tasks. In total, there were 136 survey respondents (136/526, 25.8%), including 134 surgeons who were currently practicing (124/134, 92.5%), in training (9/134, 6.7%), or recently retired (1/134, 0.7%) and 2 researchers and 1 plaster master. Nonresponse (27/136, 19.8% of respondents skipped all tasks) was explained by a problem in visual appearance when the survey was completed via computers with older Microsoft Windows editions that were not used during survey testing. No significant differences were found in the characteristics between respondents who did and did not complete the decision tasks.

Results of contextual inquiry and value specification steps of the CeHRes model are reported separately, but both were addressed during each cycle with a gradual shift in emphasis.

Contextual Inquiry

The following 7 subthemes concerning information and support needs and associated barriers and facilitators were identified in interviews with patients and focus groups with health care professionals: (1) the need for psychosocial support, (2) information on injuries and consequences of injuries, (3) information exchange between health care providers, (4) experiences of other patients, (5) workload of trauma care professionals, (6) centralized information, and (7) personalized and patient-centered care. [Table 1](#) shows details of the used themes, first-order code groups, second-order code groups, subthemes, and interview quotes.

Patients receiving psychosocial support after having polytraumatic injuries or less complex monotraumatic injuries were positive about the effects on mental health and progression in daily life activities. They experienced psychosocial support as helpful to experiences of anxiety, lack of self-efficacy, or reexperiencing traumas.

Table 1. Details of the qualitative part of the study with themes, subthemes, first-order code group examples, second-order code group examples, and interview quotes.

Themes and subthemes	First-order code group examples	Second-order code examples	Interview quote examples
Contextual inquiry			
The need for psychosocial support	<ol style="list-style-type: none"> 1. Patient information and support needs 2. Coping mechanism 	<ol style="list-style-type: none"> 1.1. Need for accessibility of biopsychosocial support 1.2. Need for reassurance 1.3. Need for humanistic approach 1.4. Need for information about affective or emotional changes 2.1. Perception of something wrong 2.2. Self-reassurance and cognitive defusion 	<p>“A few weeks after the accident, I suddenly started to cry and I could not understand why. A psychologist told me I suffered from psychological trauma. Talking about it and learning the mechanism of psychological trauma supported me in processing this trauma.” [Respondent 5]</p>
Information on injuries and consequences of injuries	<ol style="list-style-type: none"> 1. Patient information and support needs 2. Communication between patient and trauma care professional 	<ol style="list-style-type: none"> 1.1. Need for information about injury 1.2. Need for information about injury consequences 1.3. Need for information about pharmacological management 2.1. Need for support in signaling and receiving adequate responses to abnormalities or complications 	<p>“Additional information in the recovery phase would be of great benefit in comforting me in normal signals and abnormalities, what is actually normal and what is not?” [Respondent 8]</p>
Information exchange between health care providers	<ol style="list-style-type: none"> 1. Transfers to or between care settings 	<ol style="list-style-type: none"> 1.1. Transfer to other hospital department 1.2. Transfer from general practitioner to emergency department 1.3. Transfer to occupational physician 1.4. Transfer to rehabilitation center 	<p>“You need to be attentive as a patient to provide additional information. Important information was most of the times documented, but every now and then, other healthcare providers did simply not see it. This could especially be a problem while having a reduced quality of conscience due to pain medication or illness.” [Respondent 4]</p>
Experiences of other patients	<ol style="list-style-type: none"> 1. Coping mechanisms 2. Patient information and support needs 3. Potentially important pre-existing individual differences between patients 	<ol style="list-style-type: none"> 1.1. Dealing with fear of consequences 2.1. Need for social support 2.2. Need for information about affective or emotional changes 3.1. Expected individual difference in social support seeking 3.2. Expected individual difference in psychological vulnerability 	<p>“I had no need for a support group with peers, but reading about experiences of others supported me in realistic prospects and expectation management.” [Respondent 2]</p>
Workload of trauma care professionals	<ol style="list-style-type: none"> 1. Physician time restriction or work load 2. Development and implementation factors 3. Target outcome for app 	<ol style="list-style-type: none"> 1.1. Availability specialist 1.2. Asynchrony in available time between patient and physician 1.3. Change in share of routine vs nonroutine tasks 1.4. Recognition of physicians time scarcity 2.1. Reducing trauma care provider burden required for adoption 3.1. Reducing trauma care provider burden required for adoption 3.2. Reducing clinical visits 	<p>“Physiotherapists sometimes contact me in the weekends by using the communication app X (a previously introduced application) to discuss certain patients. Of course, it is my own decision to answer questions outside normal working hours, but there are already so many ways in which our tasks are being extended, that too accessible communication by patients with the doctor would be undesirable.” [Trauma surgeon 1]</p>
Centralized information	<ol style="list-style-type: none"> 1. Development and implementation factors 	<ol style="list-style-type: none"> 1.1. Preference that patients use only one app 1.2. Implementation requirement: back office or response to process patient input 	<p>“Several applications are being offered to patients, but as a patient, I would expect that all information is summarized in one tool and all communication is possible within this same tool.” [Physiotherapist 1]</p>
Personalized and patient-centered care	<ol style="list-style-type: none"> 1. Evaluations of health services 2. Patient information and support needs 	<ol style="list-style-type: none"> 1.1. Attitude toward existing services 1.2. Limited specified information 2.1. Need to personalize patient information and support to varying or unknown actual needs 	<p>“I hope, it is possible to build an application which can be self-learning to improve our standardized care.” [Trauma surgeon 2]</p>

Themes and subthemes	First-order code group examples	Second-order code examples	Interview quote examples
Value specification			
Suggestions for improvement of psychosocial and mental health	1. App attribute ideas	1.1. Hypothetical app attribute: (intelligent) monitoring and benchmarking of progress in QoL ^a and functioning (in rehabilitation phase) 1.2. Proposed app attribute: collect, model and deploy patient health data (ie, signal and responding to red flags) 1.3. Proposed app attribute: facilitated access to psychosocial help 1.4. Proposed app attribute: open field to tell your story about the event	“People should take matters into their own hands, you can assist them in monitoring psychosocial health, but they must draw their own conclusions about normality and abnormality to search for additional help on time.” [Respondent 3]
Information on injuries and consequences of injuries	1. Valuation of app attribute ideas 2. App attribute ideas 3. Target outcome for app	1.1. Attitude toward (hypothetical) technology 1.2. Perceived ease of use of (hypothetical) technology 2.1. Hypothetical app attribute: Information about common symptoms that are no reason of concern 2.2. Proposed app attribute: frequently asked questions 2.3. Hypothetical app attribute: informing about treatment procedures 3.1. Need for perspective 3.2. Recovery	“I have seen many people with functional illiteracy. Terms and language were supposed to be absolutely clear, but were not common for quite some people. That is the moment when people drop out.” [Respondent 7]
Suggestions for videos and visuals vs textual information	1. App attribute ideas	1.1. Need for guiding information or videos or photos 1.2. Proposed app attribute: possibility for sound input	“You could even add more images and graphics. A lot of people lose focus when too much text appears.” [Respondent 10]
Using surveys to detect a deviating course	1. Communication between patient and trauma care professional 2. App attribute ideas	1.1. Advantage for both patient and physician 2.1. Hypothetical app attribute: Stimulate active interaction with app to personalize content 2.2. Hypothetical app attribute: wound picture 2.3. Proposed app attribute: collecting questions or observations	“In the beginning, I was too focused on the recovery and rehabilitation of my ankle; the implications on my future life were secondary.” [Respondent 9]
Work-related information	1. Activities and participation (ICF ^b d8) 2. App attribute ideas	1.1. (Return to) work and employment (ICF: d840-d859) 1.2. Acquiring, keeping and terminating a job (ICF: d845) 2.1 Hypothetical app attribute: prompt communication with employer	“A general advice and a prospective view on return to work would be of great benefit. I do know I have to contact my employer, but when can I start working again?” [Respondent 6]

^aQoL: quality of life.

^bICF: International Classification of Functioning.

Some patients did not receive any psychosocial support, although they required psychosocial assistance. Patients did not receive support as health care professionals failed to offer it, patients did not realize the need to ask for additional support, or patients did not know where to find psychosocial support. Therefore, patients suggested standardizing the possibility of talking about emotional consequences with a social worker, spiritual caregiver, or psychologist during hospital admission. In addition, the patients suggested providing information on

where to find additional psychosocial support after hospital admission.

Extensive information on injuries and injury consequences could reduce the uncertainty of physical recovery and improve the ability to cope with limitations in daily life. Patients searched for specific information that could be used as a resource for additional information or to reread previously informed information. Participants did not demand for complicated and detailed medical information, but they appreciated receiving

basic information about the injury, treatment or treatments, complication risks, how to use painkillers, prospective on rehabilitation, and a useful prospect about the process (steps or duration) of rehabilitation. Three patients suggested using animations or short videos to discuss these topics.

Some people also missed the future prospects for returning to work. Although participants were generally satisfied with the guidance from their occupational physicians, they sometimes missed the context and information from their in-hospital health care provider. In particular, information about how injuries implied work limitations, the time span to return to work, and whether one could reasonably expect to become able to return to the *normal* working situation.

Several participants felt that communication between health care professionals from different disciplines was limited. Patients believed that most of the information was exchanged via electronic patient files and letters. The absence of face-to-face communication between health care providers may cause a loss of information. One patient with a traumatic brain injury felt that he always needed to be alert to notice mistakes during hospital admission.

Exchanging experiences with other patients often recurred during the interviews. Patients looking for leads to improve their own physical and mental health mentioned the need for like-minded experiences. Rehabilitation experiences and duration were the most commonly mentioned factors. The main goal was to obtain an impression of the illness or trauma and its subsequent consequences. Some patients had no interest in directly sharing their experiences with other patients with similar conditions. Only a few patients searched for a support group to exchange experiences, tips, and tricks, and these participants experienced benefits from a support group.

Trauma care professionals experienced several barriers and facilitators in daily trauma care. Workload was an important theme mentioned as it can act as a barrier or facilitator in introducing new web-based information tools or apps. Potentially, eHealth could relieve health care providers by supplying additional information to patients. However, the use of a mobile app or web-based application should never result in extra workload for trauma care professionals. For example, the communication capabilities of eHealth solutions should not overload professionals by shifting more work to doctors and bypass triage nurses.

Health care professionals emphasize the usefulness of an app for measurement and triage for patients potentially sustaining complications after an operation or injury. Hence, standardized but unnecessary visits might be reduced, and a shift can be made between patients who need to be seen after 2 weeks and those who can be seen after a longer period. Furthermore, 2 health care providers suggested an artificial intelligence function in which questions were automatically directed to the responsible caregiver of the specialist. This is used to balance the workload.

Different health care professionals have highlighted the importance of a centralized patient information and reporting platforms. Currently, many different platforms and tools are available, but patients and health care professionals emphasize

that a single tool would be beneficial. For example, reports or information on physical functioning from patients and physiotherapists or questionnaires could be available to the trauma care professional. This information could subsequently be used in following advice and treatment.

One health care provider summarized the use of eHealth as the ultimate goal to learn from patient outcomes to provide better targeted therapies and to adjust treatment where necessary.

Value Specification: Interviews With Patients and Professionals

On the basis of the first contextual inquiries, an existing mHealth app (patient journey) was modified and shared with patient participants during the second round interviews to facilitate value specification. These interviews revealed perceptions, attitudes, and intentions to use an eHealth application. The following five themes were identified: (1) suggestions for the improvement of psychosocial and mental health, (2) information on injuries and consequences of injuries, (3) suggestions for videos and visuals versus textual information, (4) using surveys to detect a deviating course, and (5) work-related information. [Table 1](#) presents details of the qualitative part of the study.

In addition, 60% (3/5) of participants responded positively to the proportion of self-monitoring psychosocial health outcomes. Patients and professionals agreed that data about psychosocial health, obtained by an mHealth app, could especially be used by professionals to facilitate appropriate referrals to colleagues (ie, psychologists). Furthermore, 2 patients suggested to include free text blocks in addition to structured surveys to express psychosocial difficulties. One patient suggested the use of charts for illustrating scores as a result of the questionnaires (ie, numerical pain rating score). This could indicate improvement or deterioration that requires action (receive additional advice or guidance or contact health care professionals).

Participants experienced the information in the prototype as extensive and detailed. A “read more” button for additional information was perceived as needed by 3 patients to prevent information from being too extensive. Some concerns focused on the level of education for which an app should apply. All participants agreed with the appropriateness of the current information.

Most patients missed the appropriate information on painkillers during their own recovery phase. Therefore, patients experienced information on the frequency and use of painkillers as useful. The prototype contained information on paracetamol only, but patients also requested information on additional posttraumatic pain medications such as nonsteroidal anti-inflammatory drugs and morphine.

The prototype included textual information and instruction videos, and 80% (4/5) of participants experienced this as advantageous and suggested using more videos to instruct on mobilization, exercises, or consequences of injuries. Graphical content would be easier to understand.

The prototype suggested some solutions for earlier recognition of complications and complication management. Patients especially appreciated clear information about *red flags* and

normal symptoms after treatment. Questionnaires for complication follow-up or as detectors for a deviating course or new complications were also suggested. Generic questionnaires (eg, patient-reported outcome measures or the Patient-Reported Outcomes Measurement Information System) could be used for monitoring of pain, physical function, and social or mental health. All patients stated that short surveys could be helpful in detecting problems. One patient suggested a diary function including a timeline for monitoring complaints.

In the prototype, expectations and advice on rehabilitation and recovery were divided into time spans after injury. For example, in terms of physical function, weight bearing after an operation differs between 1 and 6 weeks. All participants were positive about this feature. One patient stated that information on expectations should not be stated too early during the rehabilitation phase.

The prototype contained general advice to early contact the employer and occupational physician. Participants were generally satisfied to early contact the occupational physician and the remark that returning to work could take weeks to months, but some also questioned this general character, as working situations can differ between patients and personal advice in an app could be difficult to state. Furthermore, some stated that guidance in returning to work should be in the hands of the occupational physician, but the information on where to find additional guidance could be implemented in an app.

Value Specification: DCE With Trauma Surgeons

Results of the DCE enabled us to triangulate on parts of the interview outcomes. [Multimedia Appendix 2](#) provides a complete overview of the statistical results on the recruitment of trauma surgeons and from the predictive modeling of the surgeons' discrete choices.

Consistently across the analyses of respondents' discrete choices, patient app data access through electronic medical record integration (conditional logit odds ratio [OR] 3.33, 95% CI 2.55-4.34; $P<.001$) was weighted the highest. This was followed by the inclusion of a web viewer (the second level of the patient data access attribute; OR 2.33, 95% CI 1.64-3.31; $P<.001$); periodic self-measurements (OR 2.01, 95% CI 1.64-2.46; $P<.001$); the local adjustment of patient information (OR 1.82, 95% CI 1.42-2.33; $P<.001$); local hospital

identification (OR 1.66, 95% CI 1.31-2.10; $P<.001$); complication detection (OR 1.49, 95% CI 1.21-1.84; $P<.001$); and, lastly, the personalization of rehabilitation through artificial intelligence (OR 1.36, 95% CI 1.13-1.62; $P=.001$; [Figure 4](#)). In contrast, the estimates were negative (and $OR<1$) for price levels below €2500 (US \$2700; OR 0.82, 95% CI 0.57-1.16; $P=.26$), between €2500 (US \$2700) and €7500 (US \$8000; OR 0.66, 95% CI 0.48-0.91; $P=.01$), and between €7500 (US \$8000) and €15,000 (US \$16,000; OR 0.62, 95% CI 0.46-0.83; $P=.002$). The implications of the results for the estimate of willingness to pay, and its sensitivity to methodological choices, are visualized in [Figure 5](#). This reveals wide CIs, such that the lower bound for willingness to pay was almost 0 for artificial intelligence personalization and the upper bound for electronic medical record integration was $>€40,000$ (US \$42,761; annual). In contrast to what is expected with rational decisions, the difference between the weights for the price levels of €2500 (US \$2700) to €7500 (US \$8000) and €7500 (US \$8000) to €15,000 (US \$16,000) was relatively small compared with the absolute monetary value difference, which complicates the interpretation of willingness-to-pay results.

Furthermore, the findings suggested significant variation in preference weights across respondents. Specifically, the improvement of model fit between mixed logit model 1 (McFadden pseudo $R^2=0.27$; *good*) and conditional logit model 1 (pseudo $R^2=0.08$; not *good*) was substantial. Furthermore, models (ie, model 2) with fixed interaction effects between preference weights and respondent characteristics improved the choice predictions. The odds of selecting an app were relatively higher in surgeons with less than 10 years of work experience ($\beta=.78$, SE 0.14, $\text{Exp}[\beta]=2.20$; $P<.001$) and in those who rated the need for collective app development (very) high ($\beta=.79$, SE 0.14, $\text{Exp}[\beta]=2.19$; $P<.001$).

These weights can be used to estimate the likelihood of selecting an app under various scenarios. For example, the odds of selecting a hypothetical app with only *basic level* attributes were 1 over 5, corresponding to a probability of 0.18. Among health professionals with >10 years of experience and without a high need for collective development, the estimated probability of selecting such a basic was 8.6%. [Table 2](#) provides estimated probabilities of app selection under 4 different scenarios.

Figure 4. Weight estimates with 95% CIs for model 1 attribute levels. EMR: electronic medical record.

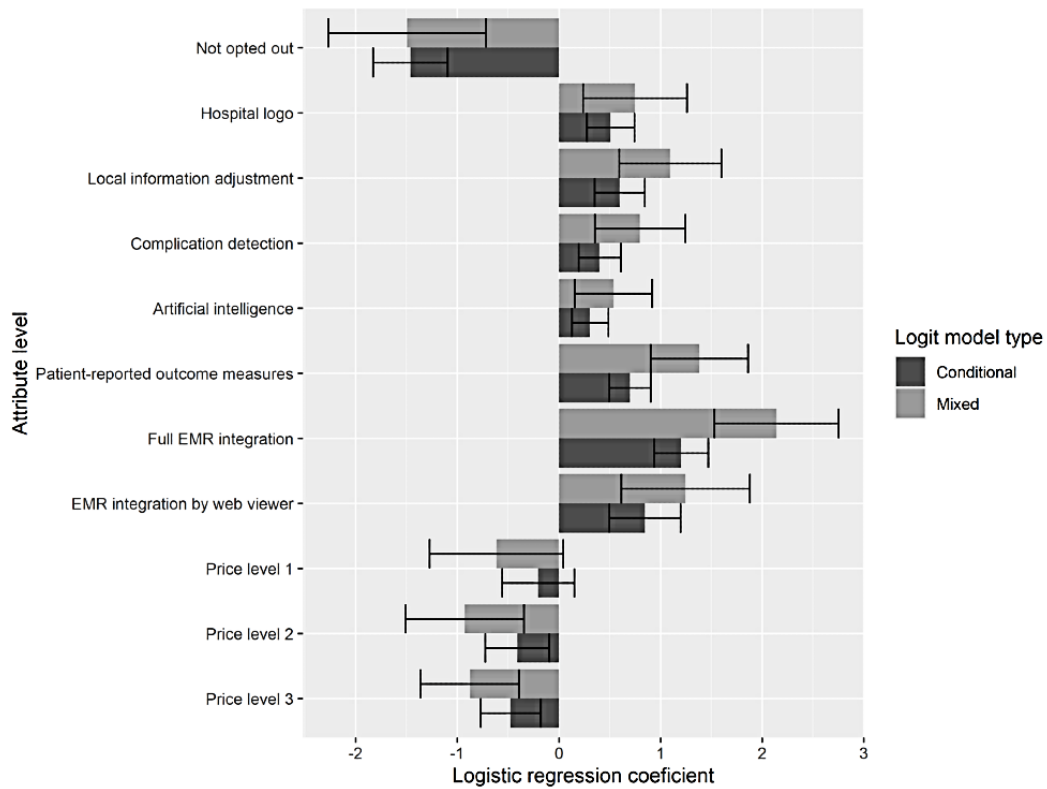


Figure 5. Willingness-to-pay estimates with 95% CIs and their sensitivity to different methodological choices. WTP: willing to pay.

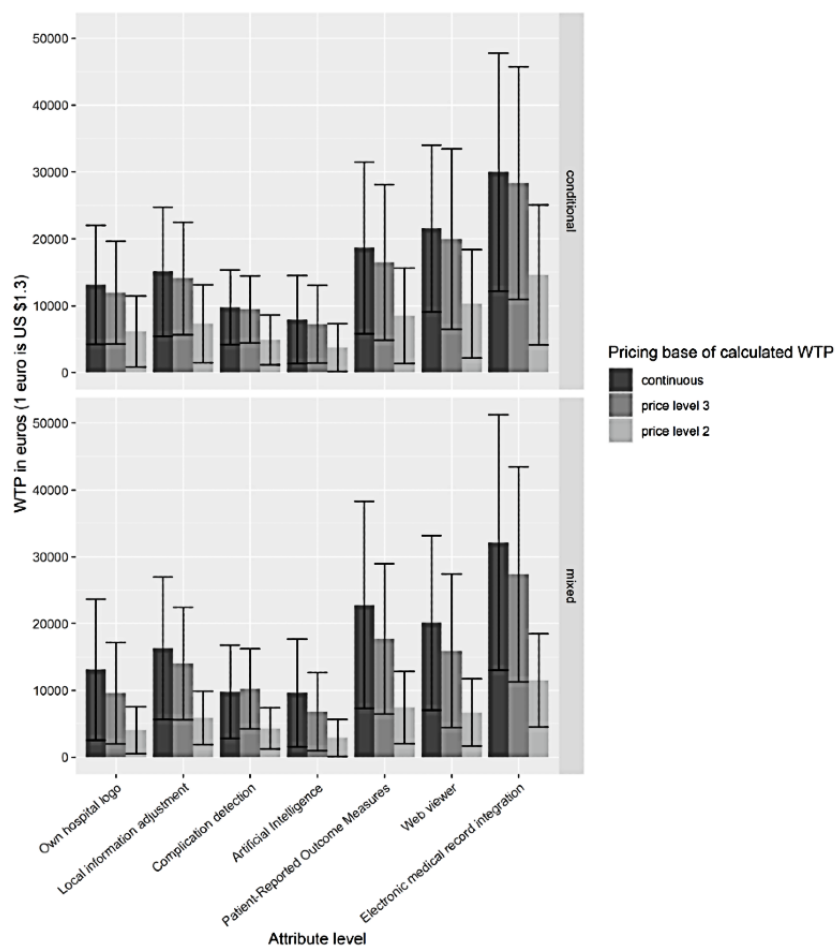


Table 2. Selection probability estimates for 4 difference scenarios of app compositions.

Scenario	Attribute levels	Estimated probability of selection, %	
		Conditional logit	Mixed logit
Basic app for free	<ul style="list-style-type: none"> • Generic logos and information • No communication options for complication detection • Standard rehabilitation plan • No patient health monitoring • No patient medical record integration of patient data access • No costs 	18	19
Best possible attributes, highest price	<ul style="list-style-type: none"> • Hospital logo and local information adjustment • Complication detection • Personalized rehabilitation plan through artificial intelligence • Patient health monitoring • Full medical record integration • €7500 (US \$8000) to €15,000 (US \$16,000) per year 	78	97
Best possible attributes except EMR ^a integration workaround	<ul style="list-style-type: none"> • As above with full medical record integration replaced with web viewer 	71	94
The two most preferred attribute levels and price level 2	<ul style="list-style-type: none"> • Generic logos and information • No communication options for complication detection • Standard rehabilitation plan • Patient health monitoring • Full medical record integration • €2500 (US \$2700) to €7500 (US \$8000) per year 	49	76

^aEMR: electronic medical record.

Discussion

Principal Findings

This study took the fundamental steps of contextual inquiry and value specification in the development of mHealth to provide injured patients with the information and support they need to recover and improve their quality of life after treatment in a (Dutch) trauma care center. We coordinated the development process among various stakeholders (eg, aligning with other initiatives with overlapping scope, addressing privacy concerns, and managing processes of integration within hospitals' existing ICT services) while maintaining a focus on end user perspectives to increase chances of future patient uptake, health care professional adoption, implementation, and scale-up across (Dutch) trauma care settings. Various formative evaluations enabled us to better understand (1) injured patients' (unmet) information and support needs, (2) barriers and facilitators in the current work of health care professionals to provide information and support provision (either with or without eHealth solutions available to them), and (3) drivers of value of an mHealth app for key user groups: patients with trauma and trauma professionals. Emerging insights informed prototyping: Selection, adjustment, and testing with an *off-the-shelf* solution with potential contextual fit. Following these steps (of the CeHRes road map) showed that eventually both trauma care professionals and patients share the need for highly accessible and valid information on how a patient's trajectory of recovery is evolving. An mHealth app including information exchange between the patient and the trauma care professional would ideally lead to a prediction model that facilitates personalized recovery trajectories.

Formative evaluation revealed important lessons. First, qualitative examinations of the (unmet) information and support needs revealed patients' need for psychosocial support and easy access to more extensive information on their injury, its consequences, and their future prospects, including return to work. Additional support and information could reduce experiences of uncertainty during physical recovery and improve the ability to cope with limitations in daily life. Second, inquiry of barriers and facilitators in the current working context of health care professionals suggested that workload is a crucial issue with regard to eHealth solutions. mHealth solutions can either act as barriers (shifting more work to doctors) or facilitators (work relief) in introducing new tools or eHealth or mHealth apps. Third, this finding explains why our DCE among trauma surgeons identified data access through electronic medical records as the most preferred attribute. In addition, albeit to a lesser extent, trauma surgeons appeared to value hospital level information adjustment and identification and personalization of rehabilitation through artificial intelligence. According to the DCE results, few trauma surgeons may be inclined to start using an mHealth app that meets none of such requirements. This seemed to be even truer the longer a surgeon is in the profession. Both the patients and professionals responded positively to using mHealth for monitoring by administering short surveys on complications, pain, physical function, and social or mental health; and receive valid feedback or prepare for inpatient consultation visits.

Comparison With Prior Work

The findings of our study add to the existing literature aimed at overcoming barriers to the successful development and implementation of eHealth initiatives that followed the CeHRes

road map [16,20,33-40]. Previous studies have addressed a wide range of eHealth initiatives such as the development of a web-based intervention for depression [37], a computerized clinical decision support system for patients with type 2 diabetes [36], an mHealth intervention for patients with prediabetes [33], and a digital training tool to support oncologists in the skill of information provision [40]. However, only one earlier study addressed a more traumatic oriented topic [38]. Although the CeHRes road map has been used in this study, numerous other frameworks are available to develop eHealth or mHealth initiatives [41]. For example, the agile science approach that includes an iterative process and focuses on flexible concepts to develop and test eHealth prototypes [42], or the persuasive system design model that focuses on influencing behavior in a positive manner [43]. Other examples include intervention mapping or the Accelerated Creation-to-Sustainment model [41]. Instead of selecting a single right development model for addressing key concerns, it is preferable to select and combine research methods to address the needs, demands, and values of end users and important stakeholders [41].

Our formative evaluations provided a relatively rare qualitative perspective on factors and domains of health-related quality of life with regard to patients' desire of well-delivered information and support during an episode of hospital trauma care. Both psychological (ie, coping, anxiety, self-efficacy, and a future prospective on the return to work) and social needs (ie, family support or a contact person in the hospital) were highlighted as essential for quality of life and progression in restoring daily life activities. Other studies have shown the influence of various consecutive (transfers between) contexts of care and support, including hospital trauma care centers, rehabilitation services [44], primary care [45], and social environments such as family [46] and work [8].

Previous studies have indicated that professionals in trauma care setting could play an important role in managing factors related to quality of life. Patient motivation, self-awareness (eg, cognitive impairment), self-efficacy (eg, managing pain or returning to work), social interaction, (work) goal setting (eg, changing occupation, following education), and eHealth or mHealth solutions could support in efficient patient guidance [7,8,47,48].

However, our qualitative look at these possibilities also highlighted that different contexts belonging to individual patients complicate the development of an eHealth app that is both personally meaningful and scalable.

Strengths and Limitations

To our knowledge, this study is the first to use the CeHRes road map for the development of a mobile delivery mode for information and support to improve experiences and multidimensional health outcomes of injured patients from trauma care settings. Using the CeHRes roadmap helped prevent common development pitfalls such as *supply drive*, *reinventing the wheel*, or a *not invented here* mindset [16,41]. During the development process, different stakeholders were involved that represented both the *demand side* (ie, patients and professionals) and the *supply side* of a *value proposition*, which is pivotal for adoption, scale-up, and maintenance of eHealth initiatives [15].

Although only patients and professionals were formally included in the study, hospital ICT services, a hospital privacy officer, and an external software developer were involved in the brainstorm sessions and prototype development. Methods used in formative evaluation during development were diverse and applied in a creative manner.

Furthermore, there were specific strengths and limitations to these formative evaluations. Our purposive sample included patients from a Dutch level I trauma center with variable injury types and complexities. This heterogeneity of patients promoted the generalizability of our findings across trauma populations.

Between the short development iterations, we embedded a quantitative method within a broader qualitative approach to triangulate on value specification from the perspective of (Dutch) trauma surgeons. Conversely, the use of qualitative data from the interviews and the focus group supported the creation of *attributes* and *levels* and compensated for the simplifying assumptions that DCEs make about complex *real-world* value attributions.

Nonetheless, several limitations to our qualitative and quantitative methods co-occurred with the challenges of aligning formative evaluation steps with those of app development and of COVID-19 restrictions. First, the foreclosure to recruit and interview patients face-to-face hindered the selection and representation of views of patients with lower levels of *eHealth literacy*. Moreover, the strategy to select a maximally heterogeneous sample of 10 patients for semistructured interviews was a choice of convenience: working toward data saturation or seeking for sampling heterogeneity in more traumatic injury dimensions were considered impractical given our goal to timely support app development with short cycles of formative evaluation. For example, a large subpopulation of patients with traumatic brain injury was represented by one patient who was also a source of insight into the uniqueness (eg, hospital boundary crossing) of the specific recovery trajectory and the associated differential needs; the needs for information related to hospital transfers and cognitive rehabilitation are minor themes within the analysis that are significant for the individual patient. Second, owing to restricted time schedules, the number of *formal* focus group meetings performed (ie, one) and the heterogeneity of roles in hospital trauma care represented in that group (trauma surgeons and physiotherapists) were smaller than desired. Third, qualitative methods mostly focused on patient perspectives, whereas quantitative methods only addressed value specification from the perspective of trauma surgeons. Thus, the principle of data triangulation was applied to part of our research scope. Full compensation for the sample size limitations of the qualitative methods would require an extended DCE sample including patients and other types of health care professionals. Consequently, it should be taken into account that our sampling strategy may have limited the generalizability of our findings regarding contextual barriers and facilitators and valued app attributes across different patient subpopulations and professional roles in trauma care. Finally, we did not perform quality assurances in the form of member checking or return interview transcripts to interview participants for possible corrections.

Implications and Recommendations

This study shows the diversity in the needs and opinions of patients and trauma care professionals regarding (digital) information provision in the current trauma care. Our findings clearly reflect the common value among patients with trauma and trauma care professionals to pursue more efficient and better-informed trauma care. A more efficient exchange of specific, valid, and standardized injury-related information and more convenient monitoring of physical, mental, and social health can be achieved by using an mHealth app. A trauma care professional who receives more information about a patient's physical, mental, and social health can better adapt his or her treatment to the individual patient. Moreover, secondary use of patient data can develop prediction models for a broad range of relevant health aspects. These models can subsequently be used to further improve shared decision-making by better-informed patients and doctors. However, this study also reminds us that such an eHealth initiative requires the broad support of patients, trauma care professionals, and other stakeholders.

The insights we obtained for the development of an eHealth prototype that fits the needs of essential end users were already used and can now be used by others to improve eHealth or mHealth apps for patients with trauma. For further development of a broadly supported eHealth app for trauma (after) care, a structural development process, for example, by means of continuing the CeHRes road map, is recommended [16].

Other authors or researchers could use this paper as inspiration for future eHealth projects in related fields of research. Herein, we recommend developers to be open to continuing iterations on each step, despite the fact that the process will move at a slow pace. The context cannot be inquired *completely* in a single

cycle with regard to all-important matters to all the stakeholders. For example, future implementation of coordinated after care supported by an mHealth solution may depend on how the sustained provision of such a service is paid for, or on legal issues regarding information exchange between patients and health care professionals of different facilities. Future studies could focus on the expansion of participants on the *supply side* and *demand side* in which, for example, outpatient nurses and clinical ward nurses could be involved in qualitative assessments. Moreover, it is strongly recommended that additional development based on contextual inquiry and, when suitable, prototype testing should emphasize on less computer or internet literate patients. When insurmountable contextual barriers are met, it is better to know them and to make the solution simpler to promote future implementation. In this regard, virtual fracture clinics provide an example to build upon.

Conclusions

This study reveals that most end users in trauma care do not just need any app or mHealth solutions. Patients, particularly those with complex injuries, require psychosocial support and easy access to more extensive information about their injury and possible journey toward recovery. Both patients with trauma and trauma care professionals strive for dependable information and, possibly, a prediction on how a patient's trajectory of recovery is evolving. Using the CeHRes road map, we were able to develop a mobile delivery mode based on the needs of both patients and trauma care professionals with basic computer skills. The formative evaluation process made it possible to iteratively adapt and improve the current prototype in an efficient way. This study could potentially serve as a starting point for future development of eHealth or mHealth initiatives within the trauma care community.

Acknowledgments

The authors would like to thank everyone who participated in the research for sharing valuable experiences and insights.

Authors' Contributions

MAPV, MCWJ, and MACdJ designed this study. MAPV, MCWJ, and TH performed data collection. MAPV, TH, and MCWJ analyzed the available data. MAPV, TH, MCWJ, MACdJ, and NN interpreted the data. MAPV and TH wrote the first version of the manuscript. TH, MAPV, MCWJ, MACdJ, NN, HPT, and KWVL revised the manuscript, and all authors approved the final version.

Conflicts of Interest

None declared.

Multimedia Appendix 1

Topic list—semistructured interview with patients.

[[DOCX File, 3618 KB - humanfactors_v9i2e35342_app1.docx](#)]

Multimedia Appendix 2

Additional details of methods of the discrete choice experiment.

[[DOCX File, 188 KB - humanfactors_v9i2e35342_app2.docx](#)]

References

1. Haagsma JA, Graetz N, Bolliger I, Naghavi M, Higashi H, Mullany EC, et al. The global burden of injury: incidence, mortality, disability-adjusted life years and time trends from the Global Burden of Disease study 2013. *Inj Prev* 2016 Mar;22(1):3-18 [FREE Full text] [doi: [10.1136/injuryprev-2015-041616](https://doi.org/10.1136/injuryprev-2015-041616)] [Medline: [26635210](https://pubmed.ncbi.nlm.nih.gov/26635210/)]
2. James SL, Theadom A, Ellenbogen RG, Bannick MS, Montjoy-Venning W, Lucchesi LR, et al. Global, regional, and national burden of traumatic brain injury and spinal cord injury, 1990–2016: a systematic analysis for the Global Burden of Disease Study 2016. *Lancet Neurol* 2019 Jan;18(1):56-87 [FREE Full text] [doi: [10.1016/S1474-4422\(18\)30415-0](https://doi.org/10.1016/S1474-4422(18)30415-0)] [Medline: [30497965](https://pubmed.ncbi.nlm.nih.gov/30497965/)]
3. Haagsma JA, Ringburg AN, van Lieshout EM, van Beeck EF, Patka P, Schipper IB, et al. Prevalence rate, predictors and long-term course of probable posttraumatic stress disorder after major trauma: a prospective cohort study. *BMC Psychiatry* 2012 Dec 27;12:236 [FREE Full text] [doi: [10.1186/1471-244X-12-236](https://doi.org/10.1186/1471-244X-12-236)] [Medline: [23270522](https://pubmed.ncbi.nlm.nih.gov/23270522/)]
4. Vardon-Bounes F, Gracia R, Abaziou T, Crognier L, Seguin T, Labaste F, et al. A study of patients' quality of life more than 5 years after trauma: a prospective follow-up. *Health Qual Life Outcomes* 2021 Jan 08;19(1):18 [FREE Full text] [doi: [10.1186/s12955-020-01652-1](https://doi.org/10.1186/s12955-020-01652-1)] [Medline: [33419450](https://pubmed.ncbi.nlm.nih.gov/33419450/)]
5. de Munter L, Polinder S, Haagsma JA, Kruithof N, van de Ree CL, Steyerberg EW, et al. Prevalence and prognostic factors for psychological distress after trauma. *Arch Phys Med Rehabil* 2020 May;101(5):877-884. [doi: [10.1016/j.apmr.2019.10.196](https://doi.org/10.1016/j.apmr.2019.10.196)] [Medline: [31874154](https://pubmed.ncbi.nlm.nih.gov/31874154/)]
6. Kruithof N, Polinder S, de Munter L, van de Ree CL, Lansink KW, de Jongh MA, BIOS-group. Health status and psychological outcomes after trauma: a prospective multicenter cohort study. *PLoS One* 2020;15(4):e0231649 [FREE Full text] [doi: [10.1371/journal.pone.0231649](https://doi.org/10.1371/journal.pone.0231649)] [Medline: [32315373](https://pubmed.ncbi.nlm.nih.gov/32315373/)]
7. Murgatroyd DF, Harris IA, Tran Y, Cameron ID, Murgatroyd D. Predictors of return to work following motor vehicle related orthopaedic trauma. *BMC Musculoskelet Disord* 2016 Apr 19;17:171 [FREE Full text] [doi: [10.1186/s12891-016-1019-6](https://doi.org/10.1186/s12891-016-1019-6)] [Medline: [27094228](https://pubmed.ncbi.nlm.nih.gov/27094228/)]
8. Hilton G, Unsworth C, Murphy G. The experience of attempting to return to work following spinal cord injury: a systematic review of the qualitative literature. *Disabil Rehabil* 2018 Jul;40(15):1745-1753. [doi: [10.1080/09638288.2017.1312566](https://doi.org/10.1080/09638288.2017.1312566)] [Medline: [28395535](https://pubmed.ncbi.nlm.nih.gov/28395535/)]
9. Thelwall C. A service evaluation after 4 year's use of the virtual fracture clinic model by a district general hospital in the South West of England. *Int J Orthop Trauma Nurs* 2021 Apr;41:100798. [doi: [10.1016/j.ijotn.2020.100798](https://doi.org/10.1016/j.ijotn.2020.100798)] [Medline: [32883628](https://pubmed.ncbi.nlm.nih.gov/32883628/)]
10. McKirdy A, Imbuldeniya AM. The clinical and cost effectiveness of a virtual fracture clinic service: an interrupted time series analysis and before-and-after comparison. *Bone Joint Res* 2017 May;6(5):259-269 [FREE Full text] [doi: [10.1302/2046-3758.65.BJR-2017-0330.R1](https://doi.org/10.1302/2046-3758.65.BJR-2017-0330.R1)] [Medline: [28473333](https://pubmed.ncbi.nlm.nih.gov/28473333/)]
11. Murphy EP, Fenelon C, Murphy RP, O'Sullivan MD, Pomeroy E, Sheehan E, et al. Are virtual fracture clinics during the COVID-19 pandemic a potential alternative for delivering fracture care? A systematic review. *Clin Orthop Relat Res* 2020 Nov;478(11):2610-2621 [FREE Full text] [doi: [10.1097/CORR.0000000000001388](https://doi.org/10.1097/CORR.0000000000001388)] [Medline: [32657810](https://pubmed.ncbi.nlm.nih.gov/32657810/)]
12. Bérubé M, Gélinas C, Feeley N, Martorella G, Côté J, Laflamme GY, et al. A hybrid web-based and in-person self-management intervention aimed at preventing acute to chronic pain transition after major lower extremity trauma: feasibility and acceptability of iPACT-E-trauma. *JMIR Form Res* 2018 Apr 30;2(1):e10323 [FREE Full text] [doi: [10.2196/10323](https://doi.org/10.2196/10323)] [Medline: [30684418](https://pubmed.ncbi.nlm.nih.gov/30684418/)]
13. Bérubé M, Gélinas C, Feeley N, Martorella G, Côté J, Laflamme GY, et al. Feasibility of a Hybrid Web-Based and In-Person Self-management Intervention Aimed at Preventing Acute to Chronic Pain Transition After Major Lower Extremity Trauma (iPACT-E-Trauma): A Pilot Randomized Controlled Trial. *Pain Med* 2019 Oct 01;20(10):2018-2032 [FREE Full text] [doi: [10.1093/pm/pnz008](https://doi.org/10.1093/pm/pnz008)] [Medline: [30840085](https://pubmed.ncbi.nlm.nih.gov/30840085/)]
14. Pfeifer A, Uddin R, Schröder-Pfeifer P, Holl F, Swoboda W, Schiltenswolf M. Mobile application-based interventions for chronic pain patients: a systematic review and meta-analysis of effectiveness. *J Clin Med* 2020 Nov 05;9(11):3557 [FREE Full text] [doi: [10.3390/jcm9113557](https://doi.org/10.3390/jcm9113557)] [Medline: [33167300](https://pubmed.ncbi.nlm.nih.gov/33167300/)]
15. Greenhalgh T, Wherton J, Papoutsi C, Lynch J, Hughes G, A'Court C, et al. Beyond adoption: a new framework for theorizing and evaluating nonadoption, abandonment, and challenges to the scale-up, spread, and sustainability of health and care technologies. *J Med Internet Res* 2017 Nov 01;19(11):e367 [FREE Full text] [doi: [10.2196/jmir.8775](https://doi.org/10.2196/jmir.8775)] [Medline: [29092808](https://pubmed.ncbi.nlm.nih.gov/29092808/)]
16. Kip H, Kelders SM, Bouman YH, van Gemert-Pijnen LJ. The importance of systematically reporting and reflecting on eHealth development: participatory development process of a virtual reality application for forensic mental health care. *J Med Internet Res* 2019 Aug 19;21(8):e12972 [FREE Full text] [doi: [10.2196/12972](https://doi.org/10.2196/12972)] [Medline: [31429415](https://pubmed.ncbi.nlm.nih.gov/31429415/)]
17. van Gemert-Pijnen JE, Nijland N, van Limburg M, Ossebaard HC, Kelders SM, Eysenbach G, et al. A holistic framework to improve the uptake and impact of eHealth technologies. *J Med Internet Res* 2011 Dec 05;13(4):e111 [FREE Full text] [doi: [10.2196/jmir.1672](https://doi.org/10.2196/jmir.1672)] [Medline: [22155738](https://pubmed.ncbi.nlm.nih.gov/22155738/)]
18. Tran V, Lam MK, Amon KL, Brunner M, Hines M, Penman M, et al. Interdisciplinary eHealth for the care of people living with traumatic brain injury: a systematic review. *Brain Inj* 2017;31(13-14):1701-1710. [doi: [10.1080/02699052.2017.1387932](https://doi.org/10.1080/02699052.2017.1387932)] [Medline: [29064300](https://pubmed.ncbi.nlm.nih.gov/29064300/)]
19. Venkatesh V, Thong JY, Xu X. Consumer acceptance and use of information technology: extending the unified theory of acceptance and use of technology. *MIS Quarterly* 2012;36(1):157. [doi: [10.2307/41410412](https://doi.org/10.2307/41410412)]

20. Daassi M, Debbabi S. Intention to reuse AR-based apps: the combined role of the sense of immersion, product presence and perceived realism. *Inf Manag* 2021 Jun;58(4):103453. [doi: [10.1016/j.im.2021.103453](https://doi.org/10.1016/j.im.2021.103453)]
21. Creswell JC, Clark VL. *Designing and Conducting Mixed Methods Research*. Thousand Oaks, California: SAGE Publications; 2006.
22. Moen K, Middelthon A. Chapter 10 - Qualitative research methods. In: *Research in Medical and Biological Sciences From Planning and Preparation to Grant Application and Publication*. Amsterdam: Academic Press; 2015.
23. Szinay D, Cameron R, Naughton F, Whitty JA, Brown J, Jones A. Understanding uptake of digital health products: methodology tutorial for a discrete choice experiment using the Bayesian efficient design. *J Med Internet Res* 2021 Oct 11;23(10):e32365 [FREE Full text] [doi: [10.2196/32365](https://doi.org/10.2196/32365)] [Medline: [34633290](https://pubmed.ncbi.nlm.nih.gov/34633290/)]
24. van Limburg M, Wentzel J, Sanderman R, van Gemert-Pijnen L. Business modeling to implement an eHealth portal for infection control: a reflection on co-creation with stakeholders. *JMIR Res Protoc* 2015 Aug 13;4(3):e104 [FREE Full text] [doi: [10.2196/resprot.4519](https://doi.org/10.2196/resprot.4519)] [Medline: [26272510](https://pubmed.ncbi.nlm.nih.gov/26272510/)]
25. *Qualitative Data Analysis A Methods Sourcebook*. Thousand Oaks, California: SAGE Publications; 2018.
26. Kostanjsek N. Use of The International Classification of Functioning, Disability and Health (ICF) as a conceptual framework and common language for disability statistics and health information systems. *BMC Public Health* 2011 May 31;11 Suppl 4:S3 [FREE Full text] [doi: [10.1186/1471-2458-11-S4-S3](https://doi.org/10.1186/1471-2458-11-S4-S3)] [Medline: [21624189](https://pubmed.ncbi.nlm.nih.gov/21624189/)]
27. Janssen EM, Hauber AB, Bridges JF. Conducting a discrete-choice experiment study following recommendations for good research practices: an application for eliciting patient preferences for diabetes treatments. *Value Health* 2018 Jan;21(1):59-68 [FREE Full text] [doi: [10.1016/j.jval.2017.07.001](https://doi.org/10.1016/j.jval.2017.07.001)] [Medline: [29304942](https://pubmed.ncbi.nlm.nih.gov/29304942/)]
28. Leigh S, Ashall-Payne L, Andrews T. Barriers and facilitators to the adoption of mobile health among health care professionals from the United Kingdom: discrete choice experiment. *JMIR Mhealth Uhealth* 2020 Jul 06;8(7):e17704 [FREE Full text] [doi: [10.2196/17704](https://doi.org/10.2196/17704)] [Medline: [32628118](https://pubmed.ncbi.nlm.nih.gov/32628118/)]
29. Campbell D, Erdem S. Including opt-out options in discrete choice experiments: issues to consider. *Patient* 2019 Feb;12(1):1-14. [doi: [10.1007/s40271-018-0324-6](https://doi.org/10.1007/s40271-018-0324-6)] [Medline: [30073482](https://pubmed.ncbi.nlm.nih.gov/30073482/)]
30. Aizaki H. Basic functions for supporting an implementation of choice experiments in R. *J Stat Softw* 2012;50(Code Snippet 2):1-24. [doi: [10.18637/jss.v050.c02](https://doi.org/10.18637/jss.v050.c02)]
31. Mulhern B, Norman R, Shah K, Bansback N, Longworth L, Viney R. How should discrete choice experiments with duration choice sets be presented for the valuation of health states? *Med Decis Making* 2018 Apr;38(3):306-318. [doi: [10.1177/0272989X17738754](https://doi.org/10.1177/0272989X17738754)] [Medline: [29084472](https://pubmed.ncbi.nlm.nih.gov/29084472/)]
32. Sarrias M, Daziano R. Multinomial logit models with continuous and discrete individual heterogeneity in : The Package. *J Stat Soft* 2017;79(2):1-46. [doi: [10.18637/jss.v079.i02](https://doi.org/10.18637/jss.v079.i02)]
33. Alshehri F, Alshaikh F. Exploring the constituent elements of a successful mobile health intervention for prediabetic patients in King Saud University Medical City Hospitals in Saudi Arabia: cross-sectional study. *JMIR Form Res* 2021 Jul 20;5(7):e22968 [FREE Full text] [doi: [10.2196/22968](https://doi.org/10.2196/22968)] [Medline: [34061762](https://pubmed.ncbi.nlm.nih.gov/34061762/)]
34. Breeman LD, Keesman M, Atsma DE, Chavannes NH, Janssen V, van Gemert-Pijnen L, BENEFIT consortium. A multi-stakeholder approach to eHealth development: promoting sustained healthy living among cardiovascular patients. *Int J Med Inform* 2021 Mar;147:104364 [FREE Full text] [doi: [10.1016/j.ijmedinf.2020.104364](https://doi.org/10.1016/j.ijmedinf.2020.104364)] [Medline: [33373949](https://pubmed.ncbi.nlm.nih.gov/33373949/)]
35. Doty JL, Brady SS, Monardez Popelka J, Rietveld L, Garcia-Huidobro D, Doty MJ, et al. Designing a mobile app to enhance parenting skills of Latinx parents: a community-based participatory approach. *JMIR Form Res* 2020 Jan 24;4(1):e12618 [FREE Full text] [doi: [10.2196/12618](https://doi.org/10.2196/12618)] [Medline: [32012034](https://pubmed.ncbi.nlm.nih.gov/32012034/)]
36. Fico G, Hernandez L, Cancela J, Dagliati A, Sacchi L, Martinez-Millana A, et al. What do healthcare professionals need to turn risk models for type 2 diabetes into usable computerized clinical decision support systems? Lessons learned from the MOSAIC project. *BMC Med Inform Decis Mak* 2019 Aug 16;19(1):163 [FREE Full text] [doi: [10.1186/s12911-019-0887-8](https://doi.org/10.1186/s12911-019-0887-8)] [Medline: [31419982](https://pubmed.ncbi.nlm.nih.gov/31419982/)]
37. Kelders SM, Pots WT, Oskam MJ, Bohlmeijer ET, van Gemert-Pijnen JE. Development of a web-based intervention for the indicated prevention of depression. *BMC Med Inform Decis Mak* 2013 Mar 20;13:26 [FREE Full text] [doi: [10.1186/1472-6947-13-26](https://doi.org/10.1186/1472-6947-13-26)] [Medline: [23425322](https://pubmed.ncbi.nlm.nih.gov/23425322/)]
38. McGrath SP, McGrath ML, Bastola D. Developing a concussion assessment mHealth app for certified Athletic Trainers. *AMIA Annu Symp Proc* 2017;2017:1282-1291 [FREE Full text] [Medline: [29854197](https://pubmed.ncbi.nlm.nih.gov/29854197/)]
39. Skjøth MM, Hansen HP, Draborg E, Pedersen CD, Lamont RF, Jørgensen JS. Informed choice for participation in down syndrome screening: development and content of a web-based decision aid. *JMIR Res Protoc* 2015 Sep 21;4(3):e113 [FREE Full text] [doi: [10.2196/resprot.4291](https://doi.org/10.2196/resprot.4291)] [Medline: [26392319](https://pubmed.ncbi.nlm.nih.gov/26392319/)]
40. Stuij SM, Drossaert CH, Labrie NH, Hulsman RL, Kersten MJ, van Dulmen S, INSTRUCT project group. Developing a digital training tool to support oncologists in the skill of information-provision: a user centred approach. *BMC Med Educ* 2020 May 01;20(1):135 [FREE Full text] [doi: [10.1186/s12909-020-1985-0](https://doi.org/10.1186/s12909-020-1985-0)] [Medline: [32357886](https://pubmed.ncbi.nlm.nih.gov/32357886/)]
41. Kip H, Keizer J, da Silva MC, Beerlage-de Jong N, Köhle N, Kelders SM. Methods for human-centered eHealth development: narrative scoping review. *J Med Internet Res* 2022 Jan 27;24(1):e31858 [FREE Full text] [doi: [10.2196/31858](https://doi.org/10.2196/31858)] [Medline: [35084359](https://pubmed.ncbi.nlm.nih.gov/35084359/)]

42. Hekler EB, Klasnja P, Riley WT, Buman MP, Huberty J, Rivera DE, et al. Agile science: creating useful products for behavior change in the real world. *Transl Behav Med* 2016 Jun;6(2):317-328 [FREE Full text] [doi: [10.1007/s13142-016-0395-7](https://doi.org/10.1007/s13142-016-0395-7)] [Medline: [27357001](https://pubmed.ncbi.nlm.nih.gov/27357001/)]
43. Asbjørnsen RA, Smedsrød ML, Solberg Nes L, Wentzel J, Varsi C, Hjelmesæth J, et al. Persuasive system design principles and behavior change techniques to stimulate motivation and adherence in electronic health interventions to support weight loss maintenance: scoping review. *J Med Internet Res* 2019 Jun 21;21(6):e14265 [FREE Full text] [doi: [10.2196/14265](https://doi.org/10.2196/14265)] [Medline: [31228174](https://pubmed.ncbi.nlm.nih.gov/31228174/)]
44. Ma VY, Chan L, Carruthers KJ. Incidence, prevalence, costs, and impact on disability of common conditions requiring rehabilitation in the United States: stroke, spinal cord injury, traumatic brain injury, multiple sclerosis, osteoarthritis, rheumatoid arthritis, limb loss, and back pain. *Arch Phys Med Rehabil* 2014 May;95(5):986-95.e1 [FREE Full text] [doi: [10.1016/j.apmr.2013.10.032](https://doi.org/10.1016/j.apmr.2013.10.032)] [Medline: [24462839](https://pubmed.ncbi.nlm.nih.gov/24462839/)]
45. Christie N, Beckett K, Earthy S, Kellezi B, Sleney J, Barnes J, et al. Seeking support after hospitalisation for injury: a nested qualitative study of the role of primary care. *Br J Gen Pract* 2016 Jan;66(642):e24-e31 [FREE Full text] [doi: [10.3399/bjgp15X688141](https://doi.org/10.3399/bjgp15X688141)] [Medline: [26639949](https://pubmed.ncbi.nlm.nih.gov/26639949/)]
46. DeHoff BA, Staten LK, Rodgers RC, Denne SC. The role of online social support in supporting and educating parents of young children with special health care needs in the United States: a scoping review. *J Med Internet Res* 2016 Dec 22;18(12):e333 [FREE Full text] [doi: [10.2196/jmir.6722](https://doi.org/10.2196/jmir.6722)] [Medline: [28007689](https://pubmed.ncbi.nlm.nih.gov/28007689/)]
47. Tamminga SJ, van Hezel S, de Boer AG, Frings-Dresen MH. Enhancing the return to work of cancer survivors: development and feasibility of the nurse-led eHealth intervention Cancer@Work. *JMIR Res Protoc* 2016 Jun 10;5(2):e118 [FREE Full text] [doi: [10.2196/resprot.5565](https://doi.org/10.2196/resprot.5565)] [Medline: [27286819](https://pubmed.ncbi.nlm.nih.gov/27286819/)]
48. Holmlund L, Guidetti S, Eriksson G, Asaba E. Return to work in the context of everyday life 7-11 years after spinal cord injury - a follow-up study. *Disabil Rehabil* 2018 Dec;40(24):2875-2883. [doi: [10.1080/09638288.2017.1362597](https://doi.org/10.1080/09638288.2017.1362597)] [Medline: [28793801](https://pubmed.ncbi.nlm.nih.gov/28793801/)]

Abbreviations

CeHRes: Center for eHealth Research and Disease Management

DCE: discrete choice experiment

ICT: information and communication technology

mHealth: mobile health

OR: odds ratio

Edited by A Kushniruk; submitted 01.12.21; peer-reviewed by A McLean, H Ayatollahi; comments to author 03.02.22; revised version received 10.05.22; accepted 14.05.22; published 20.06.22.

Please cite as:

Houwen T, Vugts MAP, Lansink KWW, Theeuwes HP, Neequaye N, Beerekamp MSH, Joosen MCW, de Jongh MAC

Developing mHealth to the Context and Valuation of Injured Patients and Professionals in Hospital Trauma Care: Qualitative and Quantitative Formative Evaluations

JMIR Hum Factors 2022;9(2):e35342

URL: <https://humanfactors.jmir.org/2022/2/e35342>

doi: [10.2196/35342](https://doi.org/10.2196/35342)

PMID: [35723928](https://pubmed.ncbi.nlm.nih.gov/35723928/)

©Thymen Houwen, Miel A P Vugts, Koen W W Lansink, Hilco P Theeuwes, Nicky Neequaye, M Susan H Beerekamp, Margot C W Joosen, Mariska A C de Jongh. Originally published in *JMIR Human Factors* (<https://humanfactors.jmir.org>), 20.06.2022. This is an open-access article distributed under the terms of the Creative Commons Attribution License (<https://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in *JMIR Human Factors*, is properly cited. The complete bibliographic information, a link to the original publication on <https://humanfactors.jmir.org>, as well as this copyright and license information must be included.

Original Paper

Implementation and Evaluation of COVIDCare@Home, a Family Medicine–Led Remote Monitoring Program for Patients With COVID-19: Multimethod Cross-sectional Study

Celia Laur^{1*}, PhD; Payal Agarwal^{1,2,3*}, MD, MSc; Kelly Thai¹, MPH; Vanessa Kishimoto¹, BSc; Shawna Kelly³, MN, NP; Kyle Liang³, MSc; R Sacha Bhatia^{1,4,5,6}, MD, MBA; Onil Bhattacharyya^{1,2}, MD, PhD; Danielle Martin^{1,2,3,7}, MD, CCFP; Geetha Mukerji^{3,5,7}, MD

¹Institute for Health System Solutions and Virtual Care (WIHV), Women's College Hospital, Toronto, ON, Canada

²Department of Family and Community Medicine, University of Toronto, Toronto, ON, Canada

³Women's College Hospital, Toronto, ON, Canada

⁴Population Health and Values Based Health Systems, Ontario Health, Toronto, ON, Canada

⁵Temerty Faculty of Medicine, University of Toronto, Toronto, ON, Canada

⁶Peter Munk Cardiac Centre, University Health Network, Toronto, ON, Canada

⁷Institute of Health Policy, Management and Evaluation, Dalla Lana School of Public Health, University of Toronto, Toronto, ON, Canada

*these authors contributed equally

Corresponding Author:

Geetha Mukerji, MD

Women's College Hospital

76 Grenville St

Toronto, ON, M5S 1B2

Canada

Phone: 1 416 323 6400

Email: Geetha.Mukerji@wchospital.ca

Abstract

Background: COVIDCare@Home (CC@H) is a multifaceted, interprofessional team-based remote monitoring program led by family medicine for patients diagnosed with COVID-19, based at Women's College Hospital (WCH), an ambulatory academic center in Toronto, Canada. CC@H offers virtual visits (phone and video) to address the clinical needs and broader social determinants of the health of patients during the acute phase of COVID-19 infection, including finding a primary care provider (PCP) and support for food insecurity.

Objective: The objective of this evaluation is to understand the implementation and quality outcomes of CC@H within the Quadruple Aim framework of patient experience, provider experience, cost, and population health.

Methods: This multimethod cross-sectional evaluation follows the Quadruple Aim framework to focus on implementation and service quality outcomes, including feasibility, adoption, safety, effectiveness, equity, and patient centeredness. These measures were explored using clinical and service utilization data, patient experience data (an online survey and a postdischarge questionnaire), provider experience data (surveys, interviews, and focus groups), and stakeholder interviews. Descriptive analysis was conducted for surveys and utilization data. Deductive analysis was conducted for interviews and focus groups, mapping to implementation and quality domains. The Ontario Marginalization Index (ON-Marg) measured the proportion of underserved patients accessing CC@H.

Results: In total, 3412 visits were conducted in the first 8 months of the program (April 8-December 8, 2020) for 616 discrete patients, including 2114 (62.0%) visits with family physician staff/residents and 149 (4.4%) visits with social workers/mental health professionals. There was a median of 5 (IQR 4) visits per patient, with a median follow-up of 7 days (IQR 27). The net promoter score was 77. In addition, 144 (23.3%) of the patients were in the most marginalized populations based on the residential postal code (as per ON-Marg). Interviews with providers and stakeholders indicated that the program continued to adapt to meet the needs of patients and the health care system.

Conclusions: Future remote monitoring should integrate support for addressing the social determinants of health and ensure patient-centered care through comprehensive care teams.

KEYWORDS

virtual care; COVID-19 pandemic; remote monitoring programs; social determinants of health; digital health; COVID-19; pandemic; health care; remote monitoring; clinical outcome; patient; health care cost; patient experience

Introduction

Remote home monitoring has dramatically expanded to manage COVID-19 in the community and avoid unnecessary hospital visits in a capacity-constrained health care system [1,2]. The ability to remotely monitor patients enables providers to escalate care at signs of deterioration, while minimizing the risk of unnecessary direct exposure of the general public, patients, and health workers to the virus [2,3]. Outcomes for remote home monitoring programs for COVID-19 are inconsistent but suggest low rates of mortality, admission rates, emergency department (ED) attendance, or reattendance [2]. Further, models of care delivery for remote monitoring vary significantly. Although many are implemented in specialist care settings [4-24], family medicine-led models may provide advantages, such as being more adaptable to meet evolving patient needs, including addressing psychosocial needs and social determinants of health within a limited capacity system [2].

To understand the impact of remote monitoring programs, evaluations of process and outcome measures are needed [25]. Greenhalgh et al [25] suggest that evaluations of COVID-19 remote monitoring programs should focus not only on the efficacy of monitoring respiratory symptoms but also on the evaluation of cost-effectiveness, patient experience, equity, sustainability, and adaptation [25]. To date, few evaluations have taken this comprehensive approach [2,16]. The Quadruple Aim framework of patient experience, provider experience, cost, and population health focuses on key process and outcome measures and is suggested as a set of principles for health system reform to be used worldwide [26].

The aim of this study was to conduct a comprehensive evaluation of the first 8 months of COVIDCare@Home (CC@H), a remote monitoring program, based at Women's College Hospital (WCH) in Toronto, Canada, that aims to address the clinical and socioeconomic needs of patients during the acute phase of COVID-19. A detailed description of the program is provided separately [27]. Lessons from this program

may be more broadly applicable to the use of remote monitoring for other acute and chronic conditions and are thus highly amenable to a primary care/family medicine approach [2]. The objectives of this evaluation are to understand the implementation and quality outcomes of CC@H within the Quadruple Aim framework of patient experience, provider experience, cost, and population health [26].

Methods

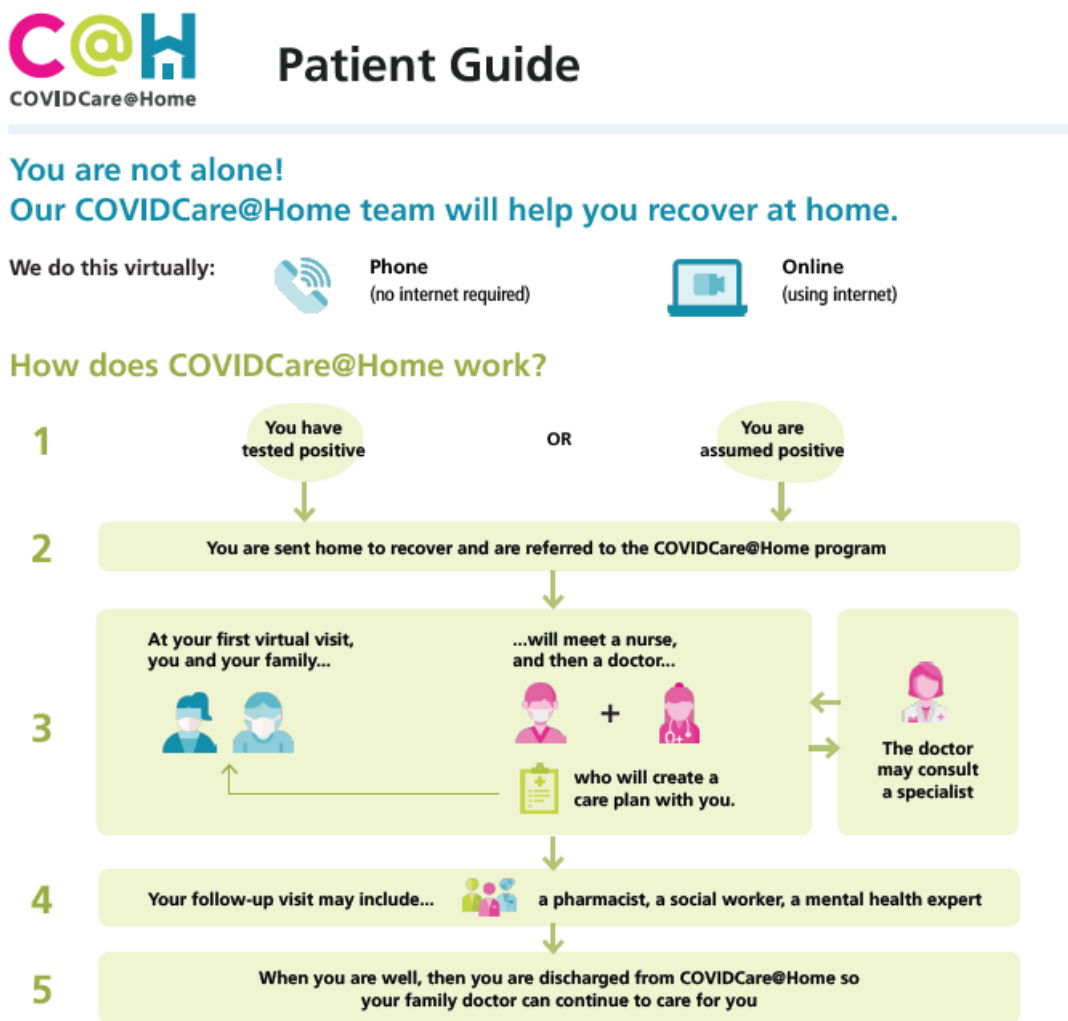
Study Design

This multimethod cross-sectional evaluation followed the Quadruple Aim framework [26] of patient experience, provider experience, cost, and population health, focusing on key process and outcome measures. Process measures included implementation outcomes of feasibility and adoption. Within patient and provider experience, the quality measures of safety, effectiveness, equity, and patient centeredness were assessed based on the National Academy of Medicine's (formerly the Institute of Medicine) domains of quality [28]. Measures were selected based on applicability to the program and feasibility of data collection. The outcome of population health included stakeholder interviews to reflect on program and health system sustainability.

Setting and Context

CC@H was launched on April 8, 2020, and is based at the WCH, an ambulatory hospital in Toronto, Ontario, Canada. Adaptive leadership and ongoing improvement cycles were used to adapt the program to meet system needs as the pandemic evolved [29]. An in-depth description of the strategies used to adapt the program and additional contextual factors are provided elsewhere [29], as are details about the model of care [27]. A patient-facing outline of the program is provided in Figure 1. In brief, the program was led by an interdisciplinary primary care team, with support from multiple specialists and allied health members. Patients were monitored at home by phone or video, sometimes with the use of a pulse oximeter.

Figure 1. Patient guide to the CC@H program. CC@H: COVIDCare@Home.



Preparing for an appointment

For a full guide on preparing for a virtual visit, go to <https://covidcareathome.ca/appointment.html>

Before your visit

Take some time to write down:

- your key symptoms and when they began
- questions you'd like to ask

The day of your visit

Leave some time (15-20 min) before your visit to get ready

- Take your measurements at rest before the call begins eg. temperature, O₂ saturation, blood pressure, heart rate
- Log in a bit early in case there are technical difficulties

Contact

Telephone:

covidcareathome.ca

Participant Recruitment

The program aims to support home-based patients with COVID-19. This study included all patients who had their first appointment in the first 8 months of the program, from April 8 to December 8, 2020. The primary referral source was the COVID-19 assessment center at the WCH. Patients were also referred by the assessment centers, ED, and acute care wards of neighboring health systems. Referred patients were excluded if they did not have access to a phone. In October 2020, to

accommodate rising case numbers, individuals aged 20-40 years who had a primary care provider (PCP) were excluded.

Outcomes

Results were organized by the Quadruple Aim framework of patient experience, provider experience, cost, and population health [26]. Process outcomes included feasibility and adoption. Within patient and provider experience, key quality measures included safety, effectiveness, equity, and patient centeredness. Data sources, samples sizes, and outcomes are summarized in Figure 2.

Figure 2. Summary of data sources and sample sizes within the Quadruple Aim framework for the CC@H evaluation. CC@H: COVIDCare@Home.

Program Feasibility & Adoption: patient clinical and utilization data (N=616), patient postdischarge survey (N=194), provider surveys (N=22), provider focus groups/interviews (N=6), and stakeholder interviews (N=8)			
Patient Experience	Provider Experience	Population Health	Cost
<ul style="list-style-type: none"> • Patient clinical and service utilization data (N=616 patients) • Patient postdischarge survey (N=194) • Online patient survey (N=14) 	<ul style="list-style-type: none"> • Provider survey (N=22; 3 time periods) • Provider focus groups/interviews (N=6) 	<ul style="list-style-type: none"> • Patient clinical and service utilization data (N=616 patients) • Patient postdischarge survey (N=194) • Provider focus groups/interviews (N=6) 	<ul style="list-style-type: none"> • Patient postdischarge survey (N=194) • Provider focus groups/interviews (N=6)
Program and Health System Sustainability: Stakeholder Interviews N=8			

Data Sources

A pragmatic approach was taken for data collection, with the aim to learn from all sources of data available. For this reason, various sources were used with variable sample sizes (Multimedia Appendix 1).

Patient Clinical and Utilization Data

Clinical information about participants was extracted from the electronic medical record (EMR) system (Epic, Epic Systems Corporation, Canada), including data entered in a standardized flowsheet (Multimedia Appendix 2). EMR data included age, sex, visit type and frequency, and length of time in the program. Flowsheet data included clinical characteristics, COVID-19-related characteristics, COVID-19 risk factors, and access to a PCP.

Patient Postdischarge Survey

Approximately 2 weeks after a patient was discharged, they received a standardized follow-up call from a nurse practitioner (NP), who asked questions verbally and entered data, including reflections and updates, into a standardized electronic flow sheet (Multimedia Appendix 3).

Online Patient Survey

An online patient evaluation survey (Multimedia Appendix 4) was developed with input from 2 lived-experience advisors. The survey was initially drafted by 2 researchers with experience in digital health evaluation, and then sent to 2 lived-experience advisors for written feedback. Once this feedback was addressed, a call was conducted between the researchers and the 2 advisors to work through each question and confirm wording and questions to add/remove, ensuring the importance and clarity of each question. Multiple scales were used in the survey as

these were selected based on the information that was most useful to the program, rather than consistency of results.

Patients who consented to be contacted at the end of the postdischarge appointment were contacted by email to complete the survey administered through Research Electronic Data Capture (REDCap), a secure, web-based software platform [30], and for their responses to be linked to the extracted clinical and utilization data using their medical record number (MRN). Surveys were sent in 2 rounds, with the first open from July 30 to August 30, 2020, and the second from November 23 to December 8, 2020. Two reminder emails were sent in each round. All participants got an opportunity to complete the survey over the phone, and a translator was available for those who preferred to answer in a language that was not English. Due to a low response rate, after the second round, patients who had provided a valid phone number received a call from a research assistant in January 2021 to complete the survey by phone.

Provider Surveys

Brief provider surveys were developed by researchers in the study team and piloted with 2 individuals who were part of the study and providers in the program. The final version was administered through Qualtrics (Multimedia Appendix 5) [31]. This voluntary survey was emailed to all CC@H providers at 3 time points (round 1: June 24, 2020; round 2: August 24, 2020; round 3: December 2, 2020), with an email reminder 1 week later. Providers included physicians, social workers/mental health professionals, nurses, NPs, and pharmacists.

Provider Interviews and Focus Groups

All providers were given the option to participate in a virtual interview or focus group. Interviews and focus groups were conducted in July 2020 by a research assistant following a semistructured guide (Multimedia Appendix 6) regarding their experience, perceived patient experience, and the impact on the

health system. Each discussion was audio-recorded then transcribed verbatim by a third party.

Stakeholder Interviews

Stakeholders, including managers and senior leadership involved in CC@H development, were recruited to participate in a semistructured, one-to-one, virtual interview. Participants were recruited by email between August 17 and October 8, 2020, with interviews conducted by a postdoctoral researcher (author CL). Interview questions ([Multimedia Appendix 7](#)) focused on health system impact, also addressing program feasibility and adoption, safety, equity, effectiveness, and patient centeredness.

Population Health: Ontario Marginalization Index (ON-Marg)

Postal code data for all participants were extracted from Epic, and the Ontario Marginalization Index (ON-Marg) [32] was calculated. ON-Marg is a data tool used to illustrate levels of marginalization across the province and combines a wide range of equity indicators based on postal code and separated by quintile [32].

Data Synthesis

For patient clinical and utilization data, descriptive analyses were conducted using R software (R Foundation for Statistical Computing), with continuous variables reported as medians (IQR) and categorical variables reported as percentages. Data were not normally distributed, so medians were used. Patient (online and postdischarge) and provider surveys were analyzed descriptively in Microsoft Excel, and the net promoter score was calculated [33]. For provider focus groups, provider interviews, and stakeholder interviews, each discussion was audio-recorded and then transcribed verbatim by a third party. Deductive content analysis was then conducted by 2 researchers (authors CL and VK) using NVivo 2020 (QSR International), mapping to implementation and service quality outcomes listed earlier within the Quadruple Aim framework. Double-coding was using to confirm results, and discrepancies were discussed with authors PA and GM. Merging of quantitative and qualitative results into the quality outcomes within the Quadruple Aim framework was an iterative process, with some data aligning with more than 1 outcome.

ON-Marg was used to provide a score to examine overall marginalization using a summated value ranging from 1 to 5, where 1 reflects low levels of marginalization and 5 reflects

high levels of marginalization. The score was used to assess the percentage of underserved patients in the CC@H program, where underserved was considered as being from the most marginalized quintile (score of 5). Individual dimensions in the score include (1) *residential instability*, area-level concentrations of people who experience high rates of family or housing instability; (2) *ethnic concentration*, high area-level concentrations of people who are recent immigrants or people belonging to a “visible minority” group, as defined by Statistics Canada; (3) *material deprivation*, closely connected to poverty and referring to inability for individuals and communities to access and attain basic material needs; and (4) *dependency*, area-level concentrations of people who do not have income from employment [32]. The ON-Marg analysis includes appointments of all statuses (completed, cancelled, etc) and repeat cases. Standard practices for calculating this score were used [32].

Ethical Considerations

This study was completed by the investigators without the influence of any commercial sponsor and was approved by the local research ethics board at the WCH (2020-0058-E).

Results

Demographics

Clinical and service utilization data were collected for all patients in the first 8 months of the program (N=616). Of the 616 patients, 337 (55%) were female, the median age was 35 (IQR 25) years, and 171 (28%) did not have a PCP. The patient postdischarge survey was conducted 2 weeks postdischarge (N=194). Of these 194 patients, 110 (57%) were female and the median age was 35 (IQR 25) years. The online patient survey was completed by N=14 patients, who had a median age of 33 (IQR 21) years and 9 (64%) of whom were female ([Table 1](#)).

Providers who completed the survey (N=22, over 3 time periods) were majorly female and represented a diverse set of clinical roles. We conducted 3 interviews and 1 focus group were conducted with CC@H clinicians (n=6, 27.3%, no physicians). Stakeholders (n=8, 36.4%; 4, 50%, female) who participated in the interviews were in managerial or leadership roles at the WCH. [Multimedia Appendix 8](#) includes the full tables of demographics and results for patients, separated by data collection tool. [Multimedia Appendix 9](#) includes the demographics and results for providers.

Table 1. Demographic and clinical characteristics of all patients (N=616) admitted to the CC@H^a program from April 8 to December 8, 2021.

Clinical and service utilization data	Patients/visits
Age (years), median 35 (IQR 25) years, n (%)	
Under 18 years of age	23 (3.7)
Over 60 years of age	85 (13.8)
Missing	508 (82.5)
Sex, n (%)	
Male	279 (45.3)
Female	337 (54.7)
Comorbidity, n (%)	
Asthma	41 (6.7)
Diabetes	36 (5.8)
Hypertension	34 (5.5)
Anxiety/depression	33 (5.4)
Other (diabetes, hypertension, etc)	45 (7.3)
Missing	427 (69.3)
Has a PCP^{b,c}, n (%)	
Yes	357 (58.0)
No	171 (27.8)
Missing	88 (14.2)
Visits (N=3412), n (%)	
Generic provider	689 (20.2)
Family physician staff/resident	2114 (62.0)
Registered nurse	439 (12.9)
Advanced nurse	2 (0.1)
Social worker/mental health professional	149 (4.4)
Pharmacist	19 (0.6)
Visits per patient, median (IQR)	5 (4)
Time from swab results to first visit, median (IQR)	3 (3)
Length of follow-up in program ^d , median (IQR)	7 (27)

^aCC@H: COVIDCare@Home.

^bPCP: primary care provider.

^cCan select more than one option.

^dTime from the first appointment to the last.

Process Outcomes: Feasibility and Adoption

Based on the utilization data (Table 1), a total of 3412 visits were conducted in the first 8 months for 616 patients, including 2114 (62.0%) visits with family physician staff/residents and 149 (4.4%) visits with a social worker/mental health professional. There was a median of 5 (IQR 4) visits per patient, with a median length of follow-up of 7 days (IQR 27). The median time from a positive swab result to the first visit was 3 days (IQR 3). All visits were conducted by phone or video, with no in-person visits.

Within the patient experience data from the patient postdischarge survey (see Multimedia Appendix 8 for full results), 177 (91.2%) of 194 patients were referred from the WCH assessment center. During the program, 39 (20.1%) patients reported they received a pulse oximeter and 14 (7.2%) received a thermometer. In addition, 60 (30.9%) patients reported receiving a referral to a social worker. Within the patient survey data (Table 2), 11 (79%) of 14 patients reported that scheduling their remote visit was easy.

From the provider survey (see Multimedia Appendix 9), most providers at each time point did not have prior experience with remote monitoring programs. All but 2 (9%) of the 22 providers

(all rounds) strongly agreed/agreed that they felt more comfortable with remote monitoring now than they did when they started the program. All but 5 (23%) providers (all rounds) strongly agreed/agreed they felt more comfortable with the technology involved in remote monitoring than when they started.

Provider interviews indicated that the initial development of the program was primarily physician driven and that involvement of nursing and allied health providers in decision-making grew as the program developed. Providers commented on the steep learning curve of rapid onboarding to a new program, adapting to delivering virtual care, using a new EMR system, challenges defining their roles and responsibilities, and getting used to the rapid decision-making needed to develop and adapt the program to meet changing patient and health

system needs. Even given these challenges, all providers recognized a strong need for the program and understood there would be challenges when developing a program so rapidly during a pandemic.

Additional facilitators mentioned in stakeholder interviews included senior leadership support, resourcing, and having regular communication between experienced clinical, operational, and technological leads. The continuous research and evaluation approach also allowed for iterations of the program, which ultimately improved care delivery. For long-term effectiveness, stakeholders valued the interdisciplinary collaboration between physicians of a variety of disciplines (ie, primary care, internal medicine), allied health professionals, academic leaders, and information management/information technology (IM/IT).

Table 2. Online patient survey data (N=14): Detailed information collected through the online patient survey focused on feasibility, adoption, safety, effectiveness, patient centeredness, and health system connection and impact. A sample of questions have been selected here, with the full results provided in [Multimedia Appendix 8](#).

Survey questions	Strongly agree, n (%)	Agree, n (%)	Neutral, n (%)	Disagree, n (%)	Strongly disagree, n (%)	N/A ^a , n (%)
Safety						
I feel my COVID-19 infection was well treated.	7 (50)	3 (21)	4 (29)	0	0	0
The health care providers had a good understanding of my medical problem(s).	7 (50)	4 (29)	2 (14)	1 (7)	0	0
I feel my care was increased when needed.	5 (36)	4 (29)	1 (7)	1 (0)	0	3 (21)
The program helped me decide if/when I needed in-person medical care.	5 (36)	4 (29)	0	0	1 (7)	4 (29)
The program helped me avoid going to the ED ^b . (Note: no patient who answered the survey went to the hospital.)	8 (57)	3 (21)	3 (21)	0	0	0
Effectiveness						
The program helped me to better manage my health and medical needs for COVID-19.	8 (57)	3 (21)	2 (14)	1 (7)	0	0
I feel I had enough time with the doctor(s).	7 (50)	5 (36)	2 (14)	0	0	0
I feel I had enough time with the other providers (ie, nurse, social worker, etc).	4 (29)	5 (36)	4 (29)	1 (7)	0	0
Patient centeredness						
I feel the care I received is in line with my goals and preferences.	7 (50)	6 (43)	0	1 (7)	0	0
This program eased my anxiety immediately after my positive COVID test.	7 (50)	4 (29)	2 (14)	1 (7)	0	0

^aN/A: not applicable.

^bED: emergency department.

Patient Experience

Equity

Of the 839 patients available in the ON-Marg data, 195 (23.2%, range 95-317, 11.3%-37.8%) were completed by patients in the most marginalized populations (marginalization score=5). Within the most marginalized, the median was 79.7% for residential instability, 74.4% for ethnic concentration, 40.4% for deprivation, and 14% for dependency.

When analyzed by visit, 564 (24.4%) of 2316 visits (range 257-831, 11.1%-35.9%) were completed by patients in the most marginalized populations. Within those most marginalized, the median by visit was 77.2% for residential instability, 75.9% for ethnic concentration, 37.6% for deprivation, and 15.8% for dependency.

Effectiveness

In the online patient survey, the net promoter score was 77 [33]. Of the 14 patients, 11 (79%) strongly agreed/agreed that the

program helped them to better manage their health and medical needs for COVID-19 and agreed that the program was useful for managing their care and treatment ([Multimedia Appendix 8](#)). In the patient postdischarge survey data, when asked about the most helpful part of the program, 69 (35.6%) of 194 patients appreciated the regular check-ins and 48 (24.7%) mentioned a positive care experience. Many patients mentioned they felt supported and reassured and that they received comprehensive, timely, and personalized care during a challenging time. Details are provided in [Multimedia Appendix 8](#).

Safety

Of the 194 patients, 10 (5.2%) in the patient postdischarge survey (see [Multimedia Appendix 8](#)) reported that since they had been diagnosed with COVID-19, they had accessed emergency services, including the ED, for COVID-19 or any other health issues. In addition, 117 (60.3%) patients felt they were discharged from CC@H at the right time and only 6 (3%) and 2 (1%) felt they were discharged too early or too late, respectively (n=70, 36%, were missing data).

Within the online patient survey data ([Multimedia Appendix 8](#)), 10 (71%) patients strongly agreed/agreed that their COVID-19 infection was well treated; 9 (64%) strongly agreed/agreed that their care was increased, when needed; and 9 (64%) strongly agreed/agreed that the program helped them decide if/when they needed in-person medical care.

Patient Centeredness

In the online patient survey data ([Multimedia Appendix 8](#)), 11 (79%) of 14 patients strongly agreed/agreed that the program eased their anxiety immediately after their positive COVID-19 test and throughout the program and 13 (93%) patients agreed/strongly agreed that their needs were addressed in the program. All but 1 (7%) patient agreed that the care they received was in line with their goals and preferences.

Provider Experience

Equity

Provider survey results (see [Multimedia Appendix 9](#)) found that all but 1 (10%) of 10 provider (round 1; 4, 40%, neutral) agreed/strongly agreed that they were able to address issues around social determinants of health for their patients. In addition, 14 (64%) of 22 providers (all 3 rounds) agreed/strongly agreed that the program was meeting the needs of underserved populations. Provider interviews demonstrated mixed opinions on whether the program initially met the needs of underserved populations, mainly focusing on the steep learning curve of navigating patients with varying immigration statuses (eg, refugees, undocumented immigrants) due to a lack of experience. Providers reported challenges in finding community resources that still offered social services throughout the COVID-19 pandemic. Resources and expertise from Crossroads Clinic, a WCH clinic specializing in refugee care, helped providers better support undocumented patients and thus improved the quality of the program [34].

Effectiveness

Interviewed providers generally agreed that CC@H was effective and met the needs of its patients. Provider interviews

also indicated provider and program flexibility were invaluable when responding to the changing environment. The primary care model and flexibility of the staffing and resources meant the team could provide comprehensive care outside the COVID-19 diagnosis to holistically support the needs of their patients. For example, due to this flexibility, a patient was able to continue within the program despite no longer presenting COVID-19 symptoms, because they required medical care but did not have access to a PCP.

We had one homeless gentleman who also had prostate cancer. He wasn't diagnosed with us, but he didn't have a family doctor. We followed him until we were able to have him see a family doctor because we got him a family doctor, but they couldn't see him for another month and a half, so we just kept monitoring him. So, there's a flexibility there to accommodate the needs of everyone. [Health care provider 1]

Safety

All but 1 (5%) provider (see [Multimedia Appendix 8](#); 3, 14%, neutral) agreed/strongly agreed that they felt supported to manage the clinical uncertainty of a new illness. All but 1 (5%; 1, 5%, neutral) provider agreed/strongly agreed they could escalate patient care, when needed. All providers (4, 18%, neutral) felt the program helped to avoid ED visits. Provider interviews discussed how frequent communication within the team about evidence and program changes was initially conducted through daily interdisciplinary group huddles to discuss patient safety and clinical issues. The primary care approach was also said to make CC@H better equipped to adapt to clinical uncertainty compared to other specialties, and thus increased the safety of the program.

In order to work in that setting [primary care], you have to be comfortable with a level of uncertainty...it's just their ability to kind of embrace the uncertainty of "you may not know the diagnosis and that's OK in family medicine." I think that's why this group of physicians was really ideally poised to take this on, because they do that every day. [Health care provider 2]

Patient Centeredness

Although patient needs varied within the program, interviewed providers generally felt that CC@H was able to improve access to medical, mental health, and social care. Specifically, they agreed that the program helped to ease patient anxiety regarding a positive COVID-19 diagnosis. Providers also commented that patients valued receiving care specific to and beyond their positive COVID-19 diagnosis, which helped to relieve their anxiety. All providers agreed/strongly agreed (3, 14%, neutral) that they could provide patient-centered care through the program, and all but 2 (9%; 3, 14%, neutral) agreed/strongly agreed that the care they could provide through the program aligned with the goals and preferences of their patients.

Population Health

Population health was explored as patient demographics (comorbidities and smoking status), access to a PCP, and receipt of community support. Utilization data indicated the most common comorbidity was asthma (n=41, 6.7%), followed by diabetes (n=36, 5.8%), hypertension (n=34, 5.5%), and anxiety/depression (n=33, 5.4%). In addition, 47 (7.6%) patients were smokers.

In the patient postdischarge survey ([Multimedia Appendix 8](#)), 32 (16.5%) of 194 patients were connected to a PCP by the program if they did not have one when entering the program. Other community support provided by CC@H included the following: 9 (4.6%) patients received food delivery, 4 (2.1%) were connected to the Red Cross, and 4 (2.1%) were connected to other types of support, such as government financial support, counseling resources, laundry, and pharmacy delivery.

In the patient postdischarge survey ([Multimedia Appendix 8](#)), when asked where they would have gone after their diagnosis if they were not involved in CC@H, 33 (17%) patients said they would go to a PCP, 17 (8.8%) said they would not have sought care, and 14 (7.2%) would have gone to the ED. From the online patient survey data ([Multimedia Appendix 8](#)), 8 (57%) patients would have gone to their PCP and 4 (29%) to the ED. When asked how many in-person visits they thought they would have had to make to a health care provider, the mean was 3 (SD 6.7, range 0-25) visits. In addition, 10 (71%) patients strongly agreed/agreed that the program could be beneficial for other patients with a lot of health issues.

Provider interviews indicated that 1 of the most valuable components of CC@H was finding PCPs for patients who did not have one. Beyond medical care, providers also reported the program was able to support patients to access groceries and medication and assist them with accessing government financial support.

Program and Health System Sustainability

CC@H stakeholder interviews focused on the sustainability of CC@H and future plans for remote monitoring. Facilitators for the sustainability of CC@H during the COVID-19 pandemic included the family medicine interdisciplinary model and having the flexibility to scale resources up and down, as needed. These facilitators were also said to support the sustainability of the health system by providing comprehensive care to patients beyond their COVID-19 diagnosis, while minimizing the risk of direct exposure to the public, patients, and health workers. Stakeholders indicated that comprehensive physician remuneration and billing codes are needed to incentivize the remote monitoring care model to improve the sustainability of remote monitoring programs in general.

Beyond the COVID-19 pandemic, stakeholders mentioned that the program could be adapted to other areas and be used to create a set of remote monitoring principles.

We have an opportunity as an organization to understand and learn from the experience that the program has with these [remote monitoring] tools...I also was always conscious of the potential for

adaptation of a program like this into very [low-resource] environments whether that be in the far north or whether it be outside our borders.
[Stakeholder 8]

Stakeholders also commented on the Quadruple Aim framework, emphasizing the importance of equitable care in the current and future iterations of the program. Comments aligned with previously mentioned experiences, with an additional point on the benefit of having diverse staff supporting diverse patients.

The [medical] residents that were engaged in the team were also very diverse, so their shared experience was helpful. In one situation we had a Black woman who really related to the fears and concerns of another Black woman who happened to have COVID, so by having a diverse group of caregivers that has also, I think, enriched the program. [Stakeholder 4]

Cost

Within the Quadruple Aim framework, low ED visits with patients can be considered proxy for cost avoidance [35]. As mentioned above, 10 (5%) patients from the post-discharge survey reported that since they had been diagnosed with COVID-19, they had accessed emergency services, including the ED, for COVID-19 or any other health issues. From the provider survey, all providers agreed/strongly agreed (4, 18%, neutral) the program has helped avoid ED visits. Interviewed providers and stakeholders perceived that CC@H prevented ED and intensive care unit (ICU) admissions.

Online patient survey data found that 3 (21%) of the 14 patients reported that they would have spent more than CA \$300 (US \$ 231.35) per visit on traveling to a health care provider (eg, parking, transit), missing work and other expenses (eg, childcare); see [Multimedia Appendix 8](#).

Provider interviews indicated that remote monitoring services have the potential to save health system costs by decreasing ED visits and hospitalizations. Stakeholder interviews discussed how crucial the implementation of the virtual billing codes is to enabling physicians to be involved in the program.

Discussion

Principal Findings

Evaluation of the CC@H remote monitoring program using the Quadruple Aim framework found the program can provide safe, effective, and patient-centered care for patients with COVID-19. With 3412 visits conducted in the first 8 months and a net promoter score of 77, the program was feasible, with care provided to a wide demographic range of patients, using primarily phone, video, and remote monitoring devices, including pulse oximeters. Our results indicate multiple benefits of remote monitoring, particularly related to patient experience. Patients highlighted the value of a continuous, hands-on touchpoint; reassurance through regular check-ins; and support in addressing the social determinants of health, including access to food, medication delivery, and a PCP.

There is also evidence that the program design enabled more equitable care, allowing the program services to reach those who were disproportionately impacted by COVID-19. In this study, 23.3% of the patients were in the most marginalized quintile patient population, suggesting an overrepresentation of patients from low-socioeconomic-status groups. This 23% is higher than the median 17.6% across the WCH.

In CC@H, 28% of the patients did not have a PCP, which is much higher than the 2019 level in Ontario, which was 9.4% (14.5% across Canada) [36]. This equity focus is especially important for COVID-19 care, as we know patients from these communities were disproportionately impacted by the disease [37-39]. The positive impact on patient-centered care and equity may reflect the comprehensive, family medicine-led, team-based design of the program.

Despite the rapid development and limited experience by providers in remote monitoring or treating patients with COVID-19, most interviews with providers and stakeholders indicated that support from leadership and team flexibility made them feel comfortable and allowed for continuous adaptations to meet evolving patient and health system needs. Almost all providers and patients felt the program helped to avoid unnecessary ED visits.

Comparison With Prior Work

Over the course of the pandemic, there has been a rapid growth in the use of remote monitoring programs to support patients and health systems. However, there is significant variation in program design and patients served (range 12-6853 patients per program), and evaluations of these programs and their impact are limited [2]. Most other programs described in the literature did not take a comprehensive, family medicine-led, team-based approach. Programs were either specialist led or focus on the use of technology as the primary mechanism for daily check-ins [4-24].

A review of COVID-19 remote monitoring programs found that few program models included support of mental health [2]. Evaluations of most programs focused on reporting adoption data (ie, number of visits) and basic clinical outcomes (eg, ED visits, deaths), with limited data presented on impact across the Quadruple Aim framework or on equity [2]. This CC@H evaluation is comprehensive, including reporting on patient, provider, and stakeholder perspectives, and assesses impact on delivery of equitable care [2]. When comparing implementation outcomes, other programs had a virtual length of stay ranging from 3.5 to 13.1 days [2] compared to CC@H's median of 7 days. Time from swab to assessment ranged from 2 to 3.7 days [2], similar to CC@H's median of 3 days. Mortality rates were also similar, ranging from 0% to 3.1% in other programs, with admission or readmission rates ranging from 0% to 29%. ED attendance or reattendance ranged from 4% to 36%, while in CC@H, it was 5%.

Most previous studies of remote monitoring programs, primarily related to cardiac disease, failed to evaluate impact of the

program on patient experience or quality of life [40-42]. Further, almost none looked specifically at the impact on social determinants of health or patients from underserved populations [40-42]. Our results suggest that future remote monitoring programs beyond the pandemic might benefit from a comprehensive team-based approach that prioritizes patient experience and support for the social determinants of health, in addition to more traditional clinical outcomes.

Limitations

This study used a pragmatic approach by leveraging the regular collection of quality improvement data, which enabled evaluation under the constraints of the pandemic; however, it also led to some issues with data quality, particularly for missing values. In data collected through EMRs, an empty response may have represented "no," "not applicable," or missing data, with no way to distinguish between these options. ED and hospital utilization data were only collected through the patient postdischarge survey based on patient reporting and may not reflect all ED visits. As providers did not always document additional services, such as support with food delivery, these values are likely lower than the actual care provided. We also were unable to collect information regarding the number of people who were excluded for not having access to a phone, as this would be an indicator of socioeconomic deprivation. Other programs at the WCH were set up to support these individuals without access to a phone. Due to our sample size, we were unable to stratify our results by waves of the pandemic.

The patient postdischarge survey data were collected by an NP as part of clinical care as it provided significant insight into the patient experience. However, data were not collected for all patients, nor were comments transcribed verbatim. This clinical approach meant data were not anonymous and were collected by someone involved in care delivery. Triangulation across multiple sources of patient experience data helped to limit the impact of all potential bias. Although this evaluation included many measures with small numbers, it is encouraging that all were pointing in the same direction, thus suggesting patient benefit. A full cost-effectiveness analysis was outside the scope of this study and deserves further exploration.

Conclusion

The CC@H remote monitoring program at the WCH is feasible and provided equitable, effective, safe, and patient-centered care during the COVID-19 pandemic. The primary care approach is thought to have facilitated comprehensive care, supporting patient needs beyond the COVID-19 diagnosis. Future remote monitoring programs should emphasize patient experience and the role of flexible, comprehensive, interdisciplinary programs that specifically address the social determinants of health. Using the Quadruple Aim framework facilitates understanding the impact of the program beyond clinical outcomes to support delivery of comprehensive, patient-centered care for all patients.

Acknowledgments

The authors would like to thank the full team of clinicians and partners who developed and continue to deliver the COVIDCare@Home (CC@H) program. Further thanks to the patients who participated in this program. The authors also wish to acknowledge the work of Shivani Chandra for conducting provider interviews and distributing online patient surveys, as well as Tanjeem Taha for sending online patient surveys and conducting patient follow-up calls. Special thanks also go to our lived experience advisors, Alies Maybee and Liana Sikharulidze, for their input to development of the online patient experience survey.

There was no direct funding provided for this study. PA is funded in part by a New Investigator Award from the Department and Community Medicine at the University of Toronto. CL was funded by the Canadian Institute for Health Research (CIHR), Health System Impact Fellowship (postdoctoral).

Data Availability

Data may be available upon request.

Authors' Contributions

PA led on all aspects of this program, with CL conducting the evaluation and drafting the manuscript. GM co-led the conception of the study and supervised the overall design and activities of the evaluation and drafting of the manuscript. KT conducted the analysis of the clinical and service utilization data. VK and CL conducted the qualitative analysis of the provider and stakeholder interviews. SK provided clinical care in the program and collected the patient postdischarge survey data. KL conducted the Ontario Marginalization Index (ON-Marg) analysis and was involved in the overall program evaluation. SB, OB, and DM provided oversight to the program development and delivery and supported manuscript conceptualization and writing. All authors read, edited, and approved the final manuscript prior to publication.

Conflicts of Interest

None declared.

Multimedia Appendix 1

COVIDCare@Home protocol.

[[PDF File \(Adobe PDF File\), 190 KB - humanfactors_v9i2e35091_app1.pdf](#)]

Multimedia Appendix 2

Electronic flowsheet for clinical and service utilization data.

[[PPTX File , 321 KB - humanfactors_v9i2e35091_app2.pptx](#)]

Multimedia Appendix 3

Patient postdischarge survey.

[[PNG File , 187 KB - humanfactors_v9i2e35091_app3.png](#)]

Multimedia Appendix 4

Online patient survey.

[[PDF File \(Adobe PDF File\), 177 KB - humanfactors_v9i2e35091_app4.pdf](#)]

Multimedia Appendix 5

Provider survey and digital consent.

[[PDF File \(Adobe PDF File\), 96 KB - humanfactors_v9i2e35091_app5.pdf](#)]

Multimedia Appendix 6

Provider interview guide.

[[PDF File \(Adobe PDF File\), 90 KB - humanfactors_v9i2e35091_app6.pdf](#)]

Multimedia Appendix 7

Stakeholder interview guide.

[[PDF File \(Adobe PDF File\), 90 KB - humanfactors_v9i2e35091_app7.pdf](#)]

Multimedia Appendix 8

Full tables of all demographics, clinical characteristics, and results collected for patients.

[PDF File (Adobe PDF File), 129 KB - [humanfactors_v9i2e35091_app8.pdf](#)]

Multimedia Appendix 9

Provider demographics and survey response focused on feasibility, adoption, safety, equity, effectiveness, patient centeredness, and cost.

[PDF File (Adobe PDF File), 62 KB - [humanfactors_v9i2e35091_app9.pdf](#)]

References

1. Monaghesh E, Hajizadeh A. The role of telehealth during COVID-19 outbreak: a systematic review based on current evidence. *BMC Public Health* 2020 Aug 01;20(1):1193 [FREE Full text] [doi: [10.1186/s12889-020-09301-4](#)] [Medline: [32738884](#)]
2. Vindrola-Padros C, Singh KE, Sidhu MS, Georghiou T, Sherlaw-Johnson C, Tomini SM, et al. Remote home monitoring (virtual wards) for confirmed or suspected COVID-19 patients: a rapid systematic review. *EClinicalMedicine* 2021 Jul;37:100965 [FREE Full text] [doi: [10.1016/j.eclinm.2021.100965](#)] [Medline: [34179736](#)]
3. Aalam AA, Hood C, Donelan C, Rutenberg A, Kane EM, Sikka N. Remote patient monitoring for ED discharges in the COVID-19 pandemic. *Emerg Med J* 2021 Mar;38(3):229-231 [FREE Full text] [doi: [10.1136/emered-2020-210022](#)] [Medline: [33472870](#)]
4. Thornton J. The "virtual wards" supporting patients with covid-19 in the community. *BMJ* 2020 Jun 04;369:m2119. [doi: [10.1136/bmj.m2119](#)] [Medline: [32499317](#)]
5. Hutchings O, Dearing C, Jagers D, Shaw M, Raffan F, Jones A, et al. Virtual health care for community management of patients with COVID-19. *medRxiv Preprint* posted online May 15, 2020. [doi: [10.1101/2020.05.11.20082396](#)]
6. Kricke G, Roemer P, Barnard C, Devin J, Henschen B, Bierman J, et al. Rapid implementation of an outpatient covid-19 monitoring program. *New Engl J Med Catal* 2020:1-12. [doi: [10.1056/CAT.20.0214](#)]
7. Annis T, Pleasants S, Hultman G, Lindemann E, Thompson J, Billecke S, et al. Rapid implementation of a COVID-19 remote patient monitoring program. *J Am Med Inform Assoc* 2020 Aug 01;27(8):1326-1330 [FREE Full text] [doi: [10.1093/jamia/ocaa097](#)] [Medline: [32392280](#)]
8. O'Keefe JB, Tong EJ, Taylor Jr TH, O'Keefe GAD, Tong DC. Initial experience in predicting the risk of hospitalization of 496 outpatients with COVID-19 using a telemedicine risk assessment tool. *medRxiv Preprint* posted online July 24, 2020 [FREE Full text] [doi: [10.1101/2020.07.21.20159384](#)]
9. Medina M, Babiuch C, Card M, Gavrilescu R, Zafirau W, Boose E, et al. Home monitoring for COVID-19. *Cleve Clin J Med* 2020 Jun 11 [FREE Full text] [doi: [10.3949/ccjm.87a.ccc028](#)] [Medline: [32409432](#)]
10. Ford D, Harvey JB, McElligott J, King K, Simpson KN, Valenta S, et al. Leveraging health system telehealth and informatics infrastructure to create a continuum of services for COVID-19 screening, testing, and treatment. *J Am Med Inform Assoc* 2020 Dec 09;27(12):1871-1877 [FREE Full text] [doi: [10.1093/jamia/ocaa157](#)] [Medline: [32602884](#)]
11. Shah S, Majmudar K, Stein A, Gupta N, Suppes S, Karamanis M, et al. Novel use of home pulse oximetry monitoring in COVID-19 patients discharged from the emergency department identifies need for hospitalization. *Acad Emerg Med* 2020 Aug;27(8):681-692 [FREE Full text] [doi: [10.1111/acem.14053](#)] [Medline: [32779828](#)]
12. Maghrabi F, Bazaz R, Wilson E, O'Reilly S, Calisti G, Richardson R, et al. S57 The development and implementation of a virtual discharge ward for patients with COVID-19 pneumonia: data on the first 300 patients. *Thorax* 2021;76(Suppl 1):A35-A36. [doi: [10.1136/thorax-2020-btsabstracts.62](#)]
13. Bell LC, Norris-Grey C, Luintel A, Bidwell G, Lanham D, Marks M, University College London Hospitals COVID response team. Implementation and evaluation of a COVID-19 rapid follow-up service for patients discharged from the emergency department. *Clin Med (Lond)* 2021 Jan;21(1):e57-e62 [FREE Full text] [doi: [10.7861/clinmed.2020-0816](#)] [Medline: [33355255](#)]
14. Lam PW, Sehgal P, Andany N, Mubareka S, Simor AE, Ozaldin O, et al. A virtual care program for outpatients diagnosed with COVID-19: a feasibility study. *CMAJ Open* 2020 May 23;8(2):E407-E413 [FREE Full text] [doi: [10.9778/cmajo.20200069](#)] [Medline: [32447283](#)]
15. Xu H, Huang S, Qiu C, Liu S, Deng J, Jiao B, et al. Monitoring and management of home-quarantined patients with COVID-19 using a WeChat-based telemedicine system: retrospective cohort study. *J Med Internet Res* 2020 Jul 02;22(7):e19514 [FREE Full text] [doi: [10.2196/19514](#)] [Medline: [32568727](#)]
16. Nunan J, Clarke D, Malakouti A, Tannetta D, Calthrop A, Hanson Xu X, et al. Triage Into the Community for COVID-19 (TICC-19) patients pathway: service evaluation of the virtual monitoring of patients with COVID pneumonia. *Acute Med J* 2020 Jan 10;19(4):183-191. [doi: [10.52964/amja.0826](#)]
17. Kodama R, Arora S, Anand S, Choudhary A, Weingarten J, Francesco N, et al. Reengineering the discharge transition process of COVID-19 patients using telemedicine, remote patient monitoring, and around-the-clock remote patient monitoring from the emergency department and inpatient units. *Telemed J E Health* 2021 Oct;27(10):1188-1193. [doi: [10.1089/tmj.2020.0459](#)] [Medline: [33320031](#)]

18. Grutters LA, Majoor KI, Mattern ESK, Hardeman JA, van Swol CFP, Vorselaars ADM. Home telemonitoring makes early hospital discharge of COVID-19 patients possible. *J Am Med Inform Assoc* 2020 Nov 01;27(11):1825-1827 [FREE Full text] [doi: [10.1093/jamia/ocaa168](https://doi.org/10.1093/jamia/ocaa168)] [Medline: [32667985](https://pubmed.ncbi.nlm.nih.gov/32667985/)]
19. Gaeta T, Chiricolo G, Mendoza C, Vaccari N, Melville L, Melniker L, et al. 124 Impact of a novel telehealth follow-up protocol for at-risk emergency department patients discharged with presumptive or confirmed COVID-19. *Ann Emerg Med* 2020 Oct;76(4):S49. [doi: [10.1016/j.annemergmed.2020.09.135](https://doi.org/10.1016/j.annemergmed.2020.09.135)]
20. Morgan A, Balachandran M, Do D, Parambath A, Chaiyachati K, Nancy M, et al. Remote Monitoring of Patients with Covid-19: design, implementation, and outcomes of the first 3,000 patients in COVID Watch. *NEJM Catal* 2020. [doi: [10.1056/CAT.20.0342](https://doi.org/10.1056/CAT.20.0342)]
21. Gordon WJ, Henderson D, DeSharone A, Fisher HN, Judge J, Levine DM, et al. Remote Patient Monitoring Program for Hospital Discharged COVID-19 Patients. *Appl Clin Inform* 2020 Oct;11(5):792-801 [FREE Full text] [doi: [10.1055/s-0040-1721039](https://doi.org/10.1055/s-0040-1721039)] [Medline: [33241547](https://pubmed.ncbi.nlm.nih.gov/33241547/)]
22. Motta LP, Silva PPF, Borguezan BM, Amaral JLMD, Milagres LG, Bóia MN, et al. An emergency system for monitoring pulse oximetry, peak expiratory flow, and body temperature of patients with COVID-19 at home: development and preliminary application. *PLoS One* 2021;16(3):e0247635 [FREE Full text] [doi: [10.1371/journal.pone.0247635](https://doi.org/10.1371/journal.pone.0247635)] [Medline: [33770093](https://pubmed.ncbi.nlm.nih.gov/33770093/)]
23. Francis NA, Stuart B, Knight M, Vancheeswaran R, Oliver C, Willcox M, et al. Predictors of clinical deterioration in patients with suspected COVID-19 managed in a 'virtual hospital' setting: a cohort study. *BMJ Open* 2021 Mar 23;11(3):e045356 [FREE Full text] [doi: [10.1136/bmjopen-2020-045356](https://doi.org/10.1136/bmjopen-2020-045356)] [Medline: [33757955](https://pubmed.ncbi.nlm.nih.gov/33757955/)]
24. Silven AV, Petrus AHJ, Villalobos-Quesada M, Dirikgil E, Oerlemans CR, Landstra CP, et al. Telemonitoring for patients with COVID-19: recommendations for design and implementation. *J Med Internet Res* 2020 Sep 02;22(9):e20953 [FREE Full text] [doi: [10.2196/20953](https://doi.org/10.2196/20953)] [Medline: [32833660](https://pubmed.ncbi.nlm.nih.gov/32833660/)]
25. Greenhalgh T, Knight M, Inda-Kim M, Fulop NJ, Leach J, Vindrola-Padros C. Remote management of covid-19 using home pulse oximetry and virtual ward support. *BMJ* 2021 Mar 25;372:n677. [doi: [10.1136/bmj.n677](https://doi.org/10.1136/bmj.n677)] [Medline: [33766809](https://pubmed.ncbi.nlm.nih.gov/33766809/)]
26. Sikka R, Morath JM, Leape L. The Quadruple Aim: care, health, cost and meaning in work. *BMJ Qual Saf* 2015 Oct;24(10):608-610. [doi: [10.1136/bmjqs-2015-004160](https://doi.org/10.1136/bmjqs-2015-004160)] [Medline: [26038586](https://pubmed.ncbi.nlm.nih.gov/26038586/)]
27. Agarwal P, Mukerji G, Laur C, Chandra S, Pimlott N, Heisey R, et al. Adoption, feasibility and safety of a family medicine-led remote monitoring program for patients with COVID-19: a descriptive study. *CMAJ Open* 2021 Apr 01;9(2):E324-E330 [FREE Full text] [doi: [10.9778/cmajo.20200174](https://doi.org/10.9778/cmajo.20200174)] [Medline: [33795222](https://pubmed.ncbi.nlm.nih.gov/33795222/)]
28. Agency for Healthcare Research and Quality. Six Domains of Health Care Quality. URL: <https://www.ahrq.gov/talkingquality/measures/six-domains.html> [accessed 2022-06-15]
29. Laur CV, Agarwal P, Mukerji G, Goulbourne E, Baranek H, Pus L, et al. Building health services in a rapidly changing landscape: lessons in adaptive leadership and pivots in a COVID-19 remote monitoring program. *J Med Internet Res* 2021 Jan 13;23(1):e25507 [FREE Full text] [doi: [10.2196/25507](https://doi.org/10.2196/25507)] [Medline: [33417588](https://pubmed.ncbi.nlm.nih.gov/33417588/)]
30. Harris PA, Taylor R, Thielke R, Payne J, Gonzalez N, Conde JG. Research electronic data capture (REDCap): a metadata-driven methodology and workflow process for providing translational research informatics support. *J Biomed Inform* 2009 Apr;42(2):377-381 [FREE Full text] [doi: [10.1016/j.jbi.2008.08.010](https://doi.org/10.1016/j.jbi.2008.08.010)] [Medline: [18929686](https://pubmed.ncbi.nlm.nih.gov/18929686/)]
31. Qualtrics. The Leading Experience Management Software. URL: <https://www.qualtrics.com/> [accessed 2022-06-15]
32. Public Health Ontario. Ontario Marginalization Index (ON-Marg). URL: <https://www.publichealthontario.ca/en/data-and-analysis/health-equity/ontario-marginalization-index> [accessed 2022-06-15]
33. NICE Systems Inc. What Is Net Promoter?. URL: <https://www.netpromoter.com/know/> [accessed 2022-06-15]
34. Women's College Hospital. Crossroads Clinic. URL: <https://www.womenscollegehospital.ca/care-programs/crossroads-clinic/> [accessed 2022-06-15]
35. Enard KR, Ganelin DM. Reducing preventable emergency department utilization and costs by using community health workers as patient navigators. *J Healthc Manag* 2013;58(6):412-427. [doi: [10.1097/00115514-201311000-00007](https://doi.org/10.1097/00115514-201311000-00007)]
36. Statistics Canada. Primary Health Care Providers, 2019. Heal Fact Sheets. URL: <https://www150.statcan.gc.ca/n1/pub/82-625-x/2020001/article/00004-eng.htm> [accessed 2022-06-15]
37. McKenzie K. Socio-demographic data collection and equity in covid-19 in Toronto. *EClinicalMedicine* 2021 Apr;34:100812. [doi: [10.1016/j.eclim.2021.100812](https://doi.org/10.1016/j.eclim.2021.100812)]
38. Kantamneni N. The impact of the COVID-19 pandemic on marginalized populations in the United States: a research agenda. *J Vocat Behav* 2020 Jun;119:103439 [FREE Full text] [doi: [10.1016/j.jvb.2020.103439](https://doi.org/10.1016/j.jvb.2020.103439)] [Medline: [32390658](https://pubmed.ncbi.nlm.nih.gov/32390658/)]
39. Garg S, Kim L, Whitaker M, O'Halloran A, Cummings C, Holstein R, et al. Hospitalization rates and characteristics of patients hospitalized with laboratory-confirmed coronavirus disease 2019: COVID-NET, 14 states, March 1–30, 2020. *Morb Mortal Wkly Rep* 2020 Apr 17;69(15):458-464. [doi: [10.15585/mmwr.mm6915e3](https://doi.org/10.15585/mmwr.mm6915e3)]
40. Lee PA, Greenfield G, Pappas Y. The impact of telehealth remote patient monitoring on glycemic control in type 2 diabetes: a systematic review and meta-analysis of systematic reviews of randomised controlled trials. *BMC Health Serv Res* 2018 Jun 26;18(1):1-10. [doi: [10.1186/s12913-018-3274-8](https://doi.org/10.1186/s12913-018-3274-8)]

41. Conway A, Inglis SC, Chang AM, Horton-Breshears M, Cleland JG, Clark RA. Not all systematic reviews are systematic: a meta-review of the quality of systematic reviews for non-invasive remote monitoring in heart failure. *J Telemed Telecare* 2013 Sep 25;19(6):326-337. [doi: [10.1177/1357633x13503427](https://doi.org/10.1177/1357633x13503427)]
42. Clark RA, Inglis SC, McAlister FA, Cleland JGF, Stewart S. Telemonitoring or structured telephone support programmes for patients with chronic heart failure: systematic review and meta-analysis. *BMJ* 2007 May 5;334(7600):942 [[FREE Full text](#)] [doi: [10.1136/bmj.39156.536968.55](https://doi.org/10.1136/bmj.39156.536968.55)] [Medline: [17426062](https://pubmed.ncbi.nlm.nih.gov/17426062/)]

Abbreviations

CC@H: COVIDCare@Home

ED: emergency department

EMR: electronic medical record

NP: nurse practitioner

ON-Marg: Ontario Marginalization Index

PCP: primary care provider

WCH: Women's College Hospital

Edited by A Kushniruk; submitted 20.11.21; peer-reviewed by D Valdes, M Mbwogge; comments to author 30.03.22; revised version received 26.04.22; accepted 01.05.22; published 28.06.22.

Please cite as:

Laur C, Agarwal P, Thai K, Kishimoto V, Kelly S, Liang K, Bhatia RS, Bhattacharyya O, Martin D, Mukerji G

Implementation and Evaluation of COVIDCare@Home, a Family Medicine-Led Remote Monitoring Program for Patients With COVID-19: Multimethod Cross-sectional Study

JMIR Hum Factors 2022;9(2):e35091

URL: <https://humanfactors.jmir.org/2022/2/e35091>

doi: [10.2196/35091](https://doi.org/10.2196/35091)

PMID: [35499974](https://pubmed.ncbi.nlm.nih.gov/35499974/)

©Celia Laur, Payal Agarwal, Kelly Thai, Vanessa Kishimoto, Shawna Kelly, Kyle Liang, R Sacha Bhatia, Onil Bhattacharyya, Danielle Martin, Geetha Mukerji. Originally published in *JMIR Human Factors* (<https://humanfactors.jmir.org>), 28.06.2022. This is an open-access article distributed under the terms of the Creative Commons Attribution License (<https://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in *JMIR Human Factors*, is properly cited. The complete bibliographic information, a link to the original publication on <https://humanfactors.jmir.org>, as well as this copyright and license information must be included.

Original Paper

Caregiver Expectations of Interfacing With Voice Assistants to Support Complex Home Care: Mixed Methods Study

Ryan Tennant^{1*}, BSc, MSc; Sana Allana^{1*}, BSc; Kate Mercer^{1,2*}, PhD; Catherine M Burns^{1*}, PEng, PhD

¹Department of Systems Design Engineering, Faculty of Engineering, University of Waterloo, Waterloo, ON, Canada

²Library, University of Waterloo, Waterloo, ON, Canada

* all authors contributed equally

Corresponding Author:

Ryan Tennant, BSc, MSc

Department of Systems Design Engineering

Faculty of Engineering

University of Waterloo

200 University Ave W

Waterloo, ON, N2L 3G1

Canada

Phone: 1 5198884567

Email: drtennan@uwaterloo.ca

Abstract

Background: Providing care in home environments is complex, and often the pressure is on caregivers to document information and ensure care continuity. Digital information management and communication technologies may support care coordination among caregivers. However, they have yet to be adopted in this context, partly because of issues with supporting long-term disease progression and caregiver anxiety. Voice assistant (VA) technology is a promising method for interfacing with digital health information that may aid in multiple aspects of being a caregiver, thereby influencing adoption. Understanding the expectations for VAs to support caregivers is fundamental to inform the practical development of this technology.

Objective: This study explored caregivers' perspectives on using VA technology to support caregiving and inform the design of future digital technologies in complex home care.

Methods: This study was part of a larger study of caregivers across North America on the design of digital health technologies to support health communication and information management in complex home care. Caregivers included parents, guardians, and hired caregivers such as personal support workers and home care nurses. Video interviews were conducted with caregivers to capture their mental models on the potential application of VAs in complex home care and were theoretically analyzed using the technology acceptance model. Interviews were followed up with Likert-scale questions exploring perspectives on other VA applications beyond participants' initial perceptions.

Results: Data were collected from 22 caregivers, and 3 themes were identified: caregivers' *perceived usefulness* of VAs in supporting documentation, care coordination, and person-centered care; caregivers' *perceived ease of use* in navigating information efficiently (they also had usability concerns with this interaction method); and caregivers' concerns, excitement, expected costs, and previous experience with VAs that influenced their *attitudes toward use*. From the Likert-scale questions, most participants (21/22, 95%) agreed that VAs should support prompted information recording and retrieval, and all participants (22/22, 100%) agreed that they should provide reminders. They also agreed that VAs should support them in an emergency (18/22, 82%)—but only for calling emergency services—and guide caregivers through tasks (21/22, 95%). However, participants were less agreeable on VAs expressing a personality (14/22, 64%)—concerned they would manipulate caregivers' perceptions—and listening ambiently to remind caregivers about their documentation (16/22, 73%). They were much less agreeable about VAs providing unprompted assistance on caregiving tasks (9/22, 41%).

Conclusions: The interviews and Likert-scale results point toward the potential for VAs to support family caregivers and hired caregivers by easing their information management and health communication at home. However, beyond information interaction, the potential impact of VA personality traits on caregivers' perceptions of the care situation and the passive collection of audio data to improve user experience through context-specific interactions are critical design considerations that should be further examined.

KEYWORDS

caregivers; children; communication; digital; home care; information management; interaction; older adult; technology acceptance; voice assistant

Introduction

Background

Although engaging in natural spoken conversation is the most common way of communicating information, humans are increasingly interacting with information through computers. The Turing test is often used to determine whether an exchange with a computer can be distinguished from that with a human, measuring the humanness of the interaction [1]. Significant research has been working toward imitating natural language conversations. However, this area has not yet been fully realized as a prominent means of human-computer interaction [2-4]. With advancements in natural language understanding and speech processing, the adoption of voice assistant (VA) technology such as Apple's Siri, Amazon's Alexa, Microsoft's Cortana, and Google's Assistant is increasing. This rise in adoption is primarily due to the ability of VAs to reduce barriers to accessing information, social attributes influencing the development of trust, and significant advancements in the technology [4-6]. Although VAs are commonly used to support everyday activities such as playing music, checking the weather, and listening to the news, emerging research explores potential health care applications [7-10].

Interacting with digital health technologies through a VA may provide a more natural, intuitive, and efficient way to engage with health information in complex home care by family members and their caregiving teams [11-13]. VAs may positively affect caregiver burnout by better supporting care coordination [14,15], where vocal recordings of health events and documentation could relieve a caregiver's documentation burdens [13]. For children with special health care needs, VAs may support autonomy to self-manage health information as they transition to adulthood [12]. At the same time, for older adults, VAs have demonstrated improvements in independent living and health maintenance [16-19].

With the increase in individuals providing home care, especially during the COVID-19 pandemic, there is significant potential for VAs to support caregivers in this context [10,13]. In 2020, approximately 1 in 5 Americans were providing home care, with an increasing number of family caregivers reporting difficulties coordinating care with other caregivers [20]. Despite the COVID-19 pandemic bringing telehealth to the forefront and the desire for integrated information technologies, there remains a lack of standardized, easy-to-use systems to support communication and coordination among caregivers in complex home care [21]. VAs may provide an interaction method that is more suitable for this health care delivery context given the atmosphere of a home environment. However, it is unclear how caregivers would expect to interact with health information using VAs, which is critical for informing their design.

Advancements in Digital Technologies for Home Care

Collaboration among caregivers is critical to ensure safety and quality care in someone's home, especially when living with complex medical conditions and health service needs [13,22-25]. Mobile apps are a promising solution to support caregiver collaboration in the home, where computer use has become ubiquitous as a technology to enhance communication and information sharing. Nursing agencies currently use mobile apps to share care updates among their teams. However, these apps are often limited to the nursing team without including the family caregiver, who ultimately develops their own information management and communication methods in the home [26]. For family caregivers supporting older adults with dementia, the design of mobile apps to meet their information and communication needs has been shown to improve caregiving confidence, depression and self-efficacy, and interaction between caregivers and health care professionals [27-29]. Mobile apps have also been shown to ease information access on the part of caregivers to scientific knowledge about their children's complex medical conditions [30].

There is increasingly more research on mobile app design, including a user-centered approach through qualitative data analysis where participants' insights and expressed needs are used to direct feature and functionality development [27,31]. These short-term deployment studies highlight the impact of the novelty factor on the interest in integrating a mobile app on the part of caregivers. However, common challenges from research on mobile app use in complex home care centers on the apps' inability to provide long-term flexibility as health conditions change or to support caregiver anxiety related to potential disease progression [30,31].

VAs may provide a way to support long-term health information management through their mode of interaction along with conversational aspects of interaction that could provide social support. Possible areas of benefit of VAs have been identified for hands-free documentation and data retrieval from electronic health records by health care professionals and for intelligent multimodal assistance by supporting telehealth use or detecting respiratory conditions [10,32]. In the context of home care, much of the current literature focuses on how older adults could interact with VAs, including medication timing and dosage reminders or encouraging physical activity [9,19]. With the rising age of our population, approximately one-third of dementia caregivers are older adults (aged >65 years) [8]. In general, older adults perceive the potential of VAs to improve their access to health information and their experiences in searching for information [33,34]. They also have concerns regarding privacy, financial burdens, and the accuracy of the information supplied. The perception of using VAs for a conversational interaction has resulted in mixed findings [33-35].

Other applications of VAs for home care have examined their use by caregivers to support older adults in aging in place and finding information, as well as for entertainment [8,35]. VAs have been designed to help caregivers manage the diet of someone diagnosed with dementia and provide guidance and personalized recommendations on nutrition, cooking, and eating behaviors [8]. Caregivers have also expressed their desire to use VAs to check in on medication events [35]. However, some of these developed systems have not been evaluated in a home care setting. Systems that have been evaluated in home care settings still experienced usability issues when integrating them into practice as the caregivers relied on paper-based tools to meet information management requirements [8,35]. There is an opportunity to use a user-centered approach to uncover aspects of VA design that should be considered to better meet the integration needs of caregivers through mixed research methods.

For caregivers of children with special health care needs, there is limited research on the potential of VAs to support health care tasks in the home. However, a spectrum of contexts for VAs has been proposed, ranging from general information retrieval to potentially prescribing therapy, medications, or other treatments [13]. VAs could also provide more autonomy to the children as they become teenagers and take more control over their health [13]. Preliminary work has shown positive attitudes toward VAs built into a medical diary app [7]. However, critical considerations and limitations preventing integration remain. For example, current limitations include access to raw health care data from mainstream vendors, Health Insurance Portability and Accountability Act compliance, the relative market demand, caregivers' social and economic status, language support, and translating current services to permit voice interaction [13].

With the potential of VAs around home care support, it is critical to better understand stakeholders' perspectives in a way that informs safe, accessible, and effective system design [10]. Few studies have explored caregivers' attitudes toward designing intelligent home-based technologies such as VAs and how they may benefit caregiving [36]. With the rise in complex home care, there is an identified need to understand the human factors influencing caregivers' perception of the usefulness and ease of use of VAs, and their attitudes toward using VAs to support technology adoption [36].

Study Objective

The objective of this study was to explore caregivers' initial perspectives on VA functionality that may influence future development and ultimately adoption of this technology using the technology acceptance model (TAM) and quantitative Likert scales. Given the collaborative nature of home care, this study included family and hired caregivers' perceptions of using VAs to interface with health information and support care coordination.

Methods

Research Design

This research is part of a larger study to identify caregivers' perspectives on information management and communication in complex home care and the design and use of VAs to support

caregivers of children with special health care needs and older adults [26]. Taking a pragmatic stance, the researchers specifically recognized that a constructivist approach to the truth must acknowledge the continuum of experiences and perspectives related to experiences, illuminating the drivers of behavior [37,38]. This paper focuses on semistructured interviews and Likert-scale question results for caregiver participants' expectations of VA functionality. The analysis was guided using the framework analysis method, which was chosen as it uses a systematic and intentionally flexible approach to analyzing multidisciplinary health and engineering data [39].

Ethics Approval

The University of Waterloo Office of Research Ethics approved this study (Office of Research Ethics 42179). All participants were interviewed via Microsoft Teams because of the COVID-19 protection measures. Informed consent was obtained verbally, and the participants received a thank-you letter for taking part in this study.

Participants and Data Collection

The research team recruited participants through home health care and caregiving agencies, social media groups, and snowball sampling. The recruitment objective was to engage participants with diverse backgrounds, ages, caregiving experiences, and experiences with VA technology in their homes. Eligible participants were either family caregivers or hired caregivers of adults or children who required complex care services in their homes in North America. In this study, complex care was defined as individuals with any combination of the following: complex chronic conditions, mental health issues, medication-related problems, and social vulnerability. A family caregiver was anyone who provided or coordinated care for a family member in their home: a parent, grandparent, guardian, spouse, child, or sibling. A hired caregiver was anyone who was paid to provide care in someone's home: a personal support worker (PSW) or a nurse that provides home care services. Participants were not required to have previous experience with VAs. Before starting the interview, the researchers explained to the participants that VAs are a technology that allows humans to interact with information on a computer system through voice and audio—the participants did not explicitly interact with a VA in this part of the research study.

In total, 2 researchers (RT and KM) conducted the interviews. First, the caregivers were asked to describe their current experiences with VAs in their daily activities. Second, the caregivers were asked to describe their initial beliefs and expectations regarding VAs to support their caregiving work domain. At the end of the interview, the participants were asked 12 Likert-scale questions about their expectations of VAs in a home care context. The participants were asked to verbally respond to each question on a 7-point scale ranging from *strongly disagree* to *strongly agree*. Microsoft Teams was used to record the interviews, and only the audio recordings were stored for transcription.

Data Analysis

The interview data on the participants' expectations of VAs in complex home care were analyzed using a theoretical thematic

process [40]. The TAM—a sociotechnical framework that posits that the adoption of a technology is driven by its capabilities and the effort required to use the technology—informed the identification of concepts and their interconnections for caregiver behaviors toward VAs in home care [41], an application context that has yet to be explored using the TAM. Although the TAM has been built upon since it was originally proposed, the fundamental framework has been successfully applied in information and communication technology in health care [42,43]. In this study, the data analysis focused on identifying the theoretical factors influencing potential usefulness, ease of use, and attitudes toward implementing VAs in home care. The usability attributes by Nielsen [44] guided the classification of the external variables influencing these 3 factors of the TAM.

First, the interviews were transcribed verbatim, and all names and identifiers were made anonymous. The research team listened to the interview recordings and read through the transcripts to familiarize themselves with the data. Core team members discussed each interview, thematically coded the data, and regularly met to discuss emerging concepts and themes. The final code list was organized into concepts and themes and presented to the entire research team for discussion and refinement. The Likert-scale results were triangulated with the participants' qualitative responses and represented graphically

while also contributing to subtheme development. These quantitative results were further broken down to visualize the expectations of participants who reported different levels of experience with VAs in their lives.

Results

Participant Demographics

There were 22 caregivers who participated in this study (Table 1). The participants were grouped by caregiver type, including family caregivers of older adults, hired caregivers of older adults, and family caregivers of children with special health care needs. The participants were recruited from various regions across Canada and the United States. The youngest participant in this study was aged 24 years, and the oldest was aged 83 years. Most of the participants identified as female (20/22, 91%), whereas 9% (2/22) identified as male. The participants' caregiving experience ranged from 4 months to 13 years. More participants reported having minimal experience with VAs (12/22, 55%) than those who did have experience with VAs (10/22, 45%). Having minimal experience was defined as understanding the concept and existence of VA technology but having little to no experience interacting with one. Being experienced was defined as owning and interacting regularly with a VA smart speaker or a VA on a mobile device.

Table 1. Participant demographics and caregiving characteristics (N=22).

Characteristics	Family caregivers of children with special health care needs (n=7), n (%)	Family caregivers of older adults (n=9), n (%)	Hired caregivers of older adults (n=6), n (%)
Age (years)			
18 to 24	0 (0)	1 (11)	0 (0)
25 to 34	2 (29)	1 (11)	1 (17)
35 to 44	5 (71)	0 (0)	2 (33)
45 to 54	0 (0)	0 (0)	1 (17)
55 to 64	0 (0)	1 (11)	1 (17)
65 to 74	0 (0)	2 (22)	1 (17)
75 to 84	0 (0)	4 (44)	0 (0)
Gender			
Female	7 (100)	8 (89)	5 (83)
Male	0 (0)	1 (11)	1 (17)
Caregiving experience (years)			
0 to 5	1 (14)	6 (67)	4 (67)
6 to 10	3 (43)	2 (22)	1 (17)
11 to 15	2 (29)	1 (11)	1 (17)
16 to 20	1 (14)	0 (0)	0 (0)
Voice assistant experience			
Minimal experience	4 (57)	5 (56)	2 (33)
Experienced	3 (43)	3 (33)	4 (67)
Unknown	0 (0)	1 (11)	0 (0)

Themes

Overview

The TAM was used to organize the qualitative findings of the participants’ initial beliefs and expectations regarding VA functionality in complex home care based on their current knowledge and experiences. There were 25 identified concepts

that were originally organized into 8 subthemes. Structured within the TAM framework (Figure 1), the similarities among participant groups supported the merging of the subthemes into 3 themes (Table 2). An additional underlying subtheme of *prior experience* was identified after further analysis of the complete data set to comprise a total of 9 subthemes.

Figure 1. Caregiving factors influencing usefulness, ease of use, and attitudes toward using a voice assistant in complex home care.

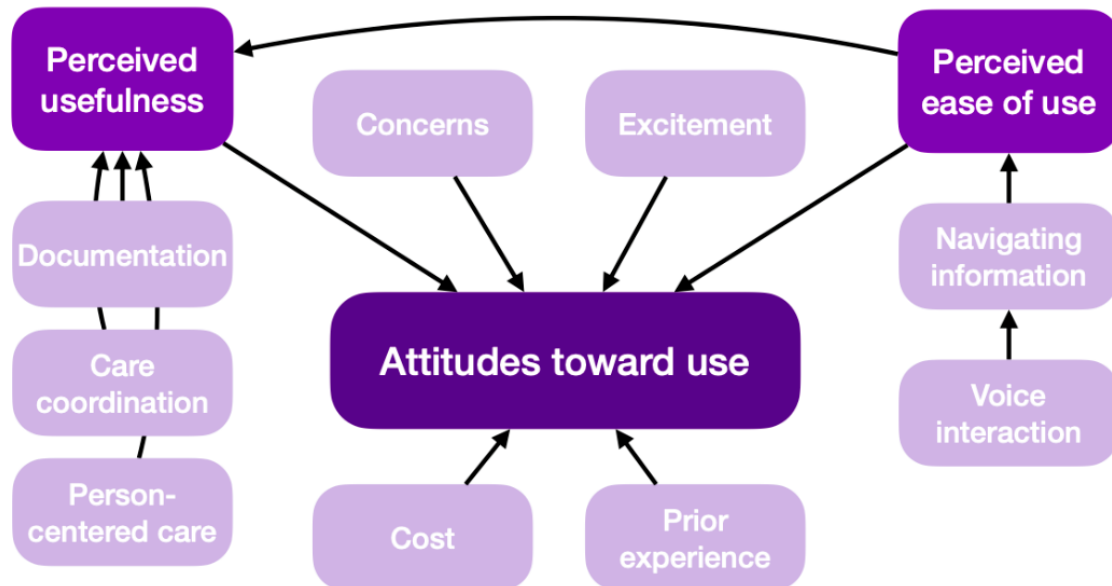


Table 2. Participants' expectations of voice assistants in complex home care (N=22).

Theme, subtheme, and concept	A ^a	B ^b	C ^c
Perceived usefulness			
Documentation			
Organizing information	✓		
Recording and retrieving information	✓	✓	✓
Care coordination			
Teaching caregivers through instructions	✓	✓	✓
Reminding caregivers	✓	✓	✓
Leaving messages for caregivers		✓	✓
Calling others	✓	✓	✓
Supporting physical tasks	✓	✓	✓
Person-centered care			
Providing autonomy for care	✓	✓	✓
Supporting mild cognitive impairment		✓	✓
Supporting medication management	✓	✓	✓
Perceived ease of use			
Navigating information efficiently			
Interacting by voice	✓	✓	
Supporting aftercare	✓		
Information retrieval		✓	
Usability concerns			
Being misunderstood or unheard	✓	✓	
Engagement by the caregiver team			✓
Challenging interfacing with computers		✓	✓
Negative influence on physical activity			✓
Attitudes toward use			
Implementation concerns			
Standard for documentation			✓
Medication management	✓		
Privacy of information	✓	✓	✓
Excitement			
Learning new technology	✓	✓	
Appreciation for voice-based technology		✓	
Excitement about home care technology			✓
Cost			
Environmental benefits			✓
Financial cost of the system	✓		

^aFamily caregivers of children with special health care needs.

^bFamily caregivers of older adults.

^cHired caregivers of older adults.

Perceived Usefulness

Despite their varied experiences with VAs, family caregiver and hired caregiver participants discussed VA design features

that would provide utility to their home care situations, which were organized into three subthemes: (1) documentation, (2) care coordination, and (3) person-centered care. First, the

participants believed that VAs would be helpful as a digital tool for managing their documentation by organizing health data and subsequently manipulating a digital record by recording and retrieving information. A participant mentioned that they would especially want to use it with a web-based notebook. They also specifically described the usefulness of maintaining documentation in the context of medication management. For example, the participants expressed that a VA could support the recording of drug reactions and the monitoring of medication adherence:

I think keeping notes, like being able to just speak out loud, and if it automatically set a date and a time for when I spoke to it with an observation that was important...if I wanted to record something about the medication. [Participant 10, experienced]

I could ask my specific question: "Alexa, did [the patient] take [their] hydromorph contin today?" [Participant 5, minimal experience]

Second, each participant population in this study discussed the VA functionalities that would affect care coordination. However, the participants had unique expectations regarding the degree to which VAs could provide coordination support. For example, the participants mentioned design functionalities that included setting reminders for medications, communicating with others, and guiding a caregiver through the steps of a medically related task:

If they got a little notice, that was like, "Hey, it's time for the medication!" I definitely think it could really be helpful. [Participant 21, minimal experience]

Certainly, managing medications, timing, and if I wanted to be reminded. [Participant 10, experienced]

The participants expressed that VAs could specifically assist with care transitions to support communication with others. For example, the participants explained that they could use the VA to leave a PSW a personal message to listen to when they arrived at their house. Some participants (3/22, 14%) also suggested that VAs would help them contact their patients or loved ones, health care professionals, or others on the caregiver team:

Well, communication with the PSWs. If I wasn't here, let's say when [my spouse]...I couldn't leave [them] alone in the latter stages. But in the earlier stages, I thought I could go off to the grocery store and leave [them]. That was up until I came home and found [them] in a delirious state and thought that was a mistake. But if I could, and I wanted to, leave instructions for a PSW.. [Participant 10, experienced]

The family caregivers of children with special health care needs detailed some of the specific contexts where a VA could support teaching their caregivers—for example, guiding caregivers through the steps involved in administering medication or operating a medical device such as a suction machine. To guide a caregiver through tasks, the participants mentioned that the caregiver could individually set a VA to provide instructions for the procedures (participant 13, minimal experience) or examples of exercises (participant 22, experienced). Although the participants who were family caregivers of children with

special health care needs in this study currently create teaching materials to support their home care, they expressed that this interaction method might positively influence the engagement of their hired caregivers with their teaching materials, improving respite care:

Taking somebody through the steps of... "This is that schedule," "This is the bottle of medication," "This is what it says," "These are the steps you go through to safely measure and administer medications." And it can be generic... "Don't touch the pills," "How to put powder in a syringe and then suck water up in it without losing all the powder." [Participant 2, minimal experience]

Family caregivers of older adults also described the use of VAs as a tool to provide instructions to caregivers where the addition of a visual representation for the steps involved in a task may improve the caregivers' capability to carry out the physical actions:

There might be able to be demonstrations of how to care for certain physical elements...Guide you...But even if it could be done, if there was a screen, if it could be done pictorial. [Participant 10, experienced]

Finally, beyond directly supporting a caregiver's tasks in the home, the participants in this study described the use of their patient or loved one interacting with the VA. They expressed that VAs could support self-care by providing autonomy in managing their medications and supporting cognitive processes and as friendly assistants to interact with during medical procedures. For example, a participant already used the reminder functionality afforded by Google Home to provide their child, who was beginning to take more responsibility for their care, with more autonomy in taking their medications:

We had the medication set up all around, kind of in [their downstairs] apartment. So, we set it up, you know, "set a reminder for [them] to take the pills on top of your white freezer with the Green Cup at 8:00." [Participant 8, experienced]

For adults who may have mild cognitive impairment or physical disabilities, the participants expressed that VAs could support their autonomy through reminders about their care. For example, a caregiver mentioned that VAs could help an older adult through reminders, specifically when to expect their hired care to arrive, without finding the information physically:

If you could have said things like, "Siri, what time does my home care person arrive?" And if it could have given the appointment time to [them] verbally [they] wouldn't have had to search through papers. [Participant 13, minimal experience]

The participants mentioned that caregivers could interact with a VA to check whether a patient or loved one had taken their medication. A care receiver could check their complete medication history using a command. The participants also discussed the importance of supporting cognitive processes to keep older adults oriented with their environment and assist with medication management through verbal cues:

Having a verbal cue for the person to take their medication but as a backup. Seeing if it has been done. [Participant 5, minimal experience]

I'm beginning to think...something to remind you when and how often you've taken your medication would be good. [Participant 14, minimal experience]

In the context of interacting with VAs during a medical procedure, a participant described that a VA could interact with their child to keep them calm while they changed their tracheostomy:

[My child] could like use it to talk [them] through a medical procedure, and that might calm [their] anxiety down a bit...And just like in a kid-friendly way...that would be cool to have in-home. [Participant 1, minimal experience]

Perceived Ease of Use

In this study, participants with varied experiences using a VA commented explicitly on the ease of using a VA in a home care context, organized into two subthemes: (1) navigating information efficiently and (2) usability concerns. First, the participants mentioned that VAs would ease their documentation. They also expressed that interacting by voice would facilitate recording and retrieving information as it only takes as long as they need to talk. A participant also commented that a voice-based system could instantaneously give information compared with a paper notebook.

The participants described the affordance of multitasking that a VA could provide. They expressed that, while working on a task, they could speak to the system and have health information documented directly during that moment. The ease of recording by voice may reduce the burden of physically writing information on paper; however, the participants still desired to obtain a physical copy of the data if needed:

Sometimes I'm in the middle of doing something else...and I need to remember this thing. But if I stop what I'm doing, then...maybe it's not that simple to just stop what I'm doing. Or if I wait until the end, I'm going to forget because I just don't have a very good memory... [Participant 3, experienced]

The ease of record keeping by voice could also support a caregiver's capacity to perform aftercare. For example, a participant mentioned that, if their child were having a seizure, they would be able to physically care for them while maintaining accurate documentation of the event:

If my [child's] in the middle of a seizure: "Siri, note that [they] had started a seizure at this time," "Siri, note that [they] stopped," so I'm not having to wait for [them] to get done and try to remember all the time. [Participant 16, minimal experience]

Second, despite the design functionality of VAs that would ease documentation, there were essential concerns regarding this method of interacting with health information. A caregiver (participant 3, experienced) mentioned that using a VA may not be a more straightforward method for managing their child's health information. However, they first expressed the need to

integrate the technology into their routine to determine whether it would be a valuable alternative to other technologies, processes, or practices. There were also concerns about their voice commands being accurately understood by a VA, which may lead to a problematic interaction:

[Siri] just...it wouldn't register what I was saying...if I have that [for home care], is it going to even register what I'm saying? [Participant 16, minimal experience]

A participant was strongly opposed to interacting with VAs in complex home care. Their perceived trust in VA technology, hesitations about information privacy, and the accuracy of recording information by voice negatively influenced their perceived ease of use. Although the participants identified the need for all members of the caregiver team to be comfortable interacting with the VA, conflicting beliefs about the ease of record keeping using a VA might negatively influence care coordination:

I don't think it's a good idea; I don't like that idea. Things can get messed up. You know, certain things could be left out. I mean, it's always glitches with computers, and they frustrate me all the time. [Participant 4, experienced]

Another participant mentioned an essential caveat for technology such as VAs being easy to use. Although they believed that VAs might support individuals with mild cognitive impairment, their concern was that this might negatively influence their physical activity as other technologies have done in the past:

I must admit I have real reservations about them; the more electronics do for us physically...The two things that, for health for seniors and keeping them in their home, they have to have mobility, and I mean I can see it supporting cognition. Things to keep them in their home longer. It's like the remote on the TV. That getting up and moving to turn on the TV used to be sometimes the only activity those seniors see. So, I'm not sure it's necessarily a good thing in that respect. [Participant 13, minimal experience]

Attitudes Toward Use

The participants in this study were excited to think about what they could do with the technology. Despite their varied experiences, they initially expressed excitement about integrating digital home care solutions and their willingness to learn a new technology that could support their caregiving tasks (4/22, 18%). The external factors influencing the participants' behavioral intention to use VAs in complex home care were organized into four subthemes: (1) excitement, (2) implementation concerns, (3) cost, and (4) prior experience. It is important to note that, although the participants did not explicitly comment on how their previous experience with VAs influenced their attitudes, the fourth subtheme was developed and explored through a deeper analysis of the Likert-scale results in the subsequent section.

Concerns about using VAs were grounded in the current methods the participants used to document health information in the home. As a first example, a participant (participant 4, experienced) explained that health information should not be

obtained from a VA but should come directly from the patient or other caregivers. Another participant mentioned that using VAs for medication management may not be as accurate as their current system for tracking their child’s complex medication regimen, which currently provided a physical cue for measuring adherence:

For example, remembering to take [their] meds. I don’t know that I would use [a voice assistant] for that, and the reason being...you can forget to tell it that [you] took it, but...my little pillbox doesn’t lie. So, if it’s in there, I know you didn’t take it. [There’s] no “I just forgot to tell it,” “I actually did take it,” kind of thing. [Participant 3, experienced]

Privacy of information was also an essential concern for the participants. In one situation, the family members of a participant (participant 12, minimal experience) influenced them not to purchase a VA based on the perception that they will always listen to what is going on in their homes. Another participant (participant 4, experienced) further expressed concerns about others accessing someone’s health information stored on VAs.

Finally, although the financial cost was an initial concern mentioned by a participant in this study, a hired caregiver also

noted the cost of their current documentation methods to our environment and how the use of VAs could support the reduction of that cost:

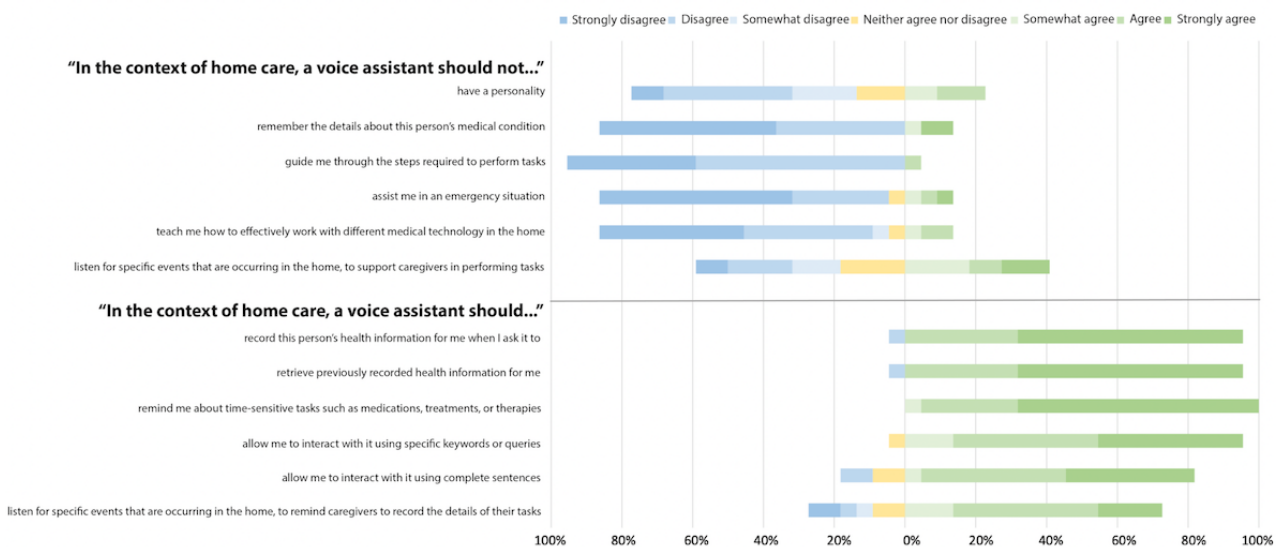
When it’s paper-based, it’s basically really a big waste...of paper. So, at least if you’re just using Alexa or a voice assistant...it would be at least...let’s say...kinder to nature...If we’re looking at [my client’s] records of [their] things, whenever we try to record the chart, we basically have a load thick of this paper. [Participant 20, experienced]

Likert-Scale Results

Overview

The caregivers’ initial mental models on using VAs were analyzed to provide insights into how design decisions may affect the successful integration of VAs into complex home care. The Likert-scale questions were used after the interview to prompt additional discussion on the potential features of a VA for complex home care. The Likert-scale questions captured the participants’ initial perspectives on specific design features for VAs in complex home care while exploring their opinions on potential functionality beyond their current mental models. We represented these results graphically to visualize aspects of VA expectations (Figure 2).

Figure 2. Caregiver expectations of voice assistants in complex home care (N=22).



Overall, most participants agreed that a VA should record someone’s health information when they request it (21/22, 95%) and retrieve previously documented information (21/22, 95%). They also agreed that a VA should remember the details of someone’s medical condition (19/22, 86%), with the requirement that the data not be stored in a publicly accessible database:

As long as there’s privacy, I think it should. It should be able to retain it. If I came in as a home care nurse or PSW, even as a family member, and I say, “When did this happen?” I don’t have to go back through my notes. My machine can testify who did the treatment last. I mean, that would be very

helpful...Anything that records, and I don’t have to chart, I’m on board! [Participant 13, minimal experience]

All the study participants (22/22, 100%) agreed that VAs should remind them about time-sensitive tasks such as medications, treatments, or therapies. With respect to interaction preferences with VAs, the participants often expressed their desire to have the option to speak using specific keywords (21/22, 95%) and complete sentences (18/22, 82%).

For more dynamic interactions, most participants agreed that VAs should guide them through the steps required to perform tasks (21/22, 95%), teach them how to use different medical

technologies in their homes (18/22, 82%), and support them in an emergency (18/22, 82%). However, they were relatively less agreeable about VAs having a personality (14/22, 64%). The participants were also less agreeable about VAs listening for a particular activity in the home to remind caregivers to record the details of their tasks (16/22, 73%), where more experienced participants (3/10, 30%) disagreed that VAs should listen in this context compared with minimally experienced participants (1/11, 9%). The remaining participants (2/22, 9%) were unsure. Fewer participants (9/22, 41%) felt that a VA should listen for

a specific activity in the home to support caregivers in performing their tasks, whereas more experienced participants (6/10, 60%) agreed that it should not compared with minimally experienced participants (2/11, 18%). The remaining participants (4/22, 18%) were unsure. The contrasting perspectives on VA personality, support in an emergency, teaching or guiding caregivers, and listening to activity in the home are further visualized in Figures 3 and 4 and described in the following sections.

Figure 3. Caregiver expectations of voice assistants in complex home care—experienced (n=10).

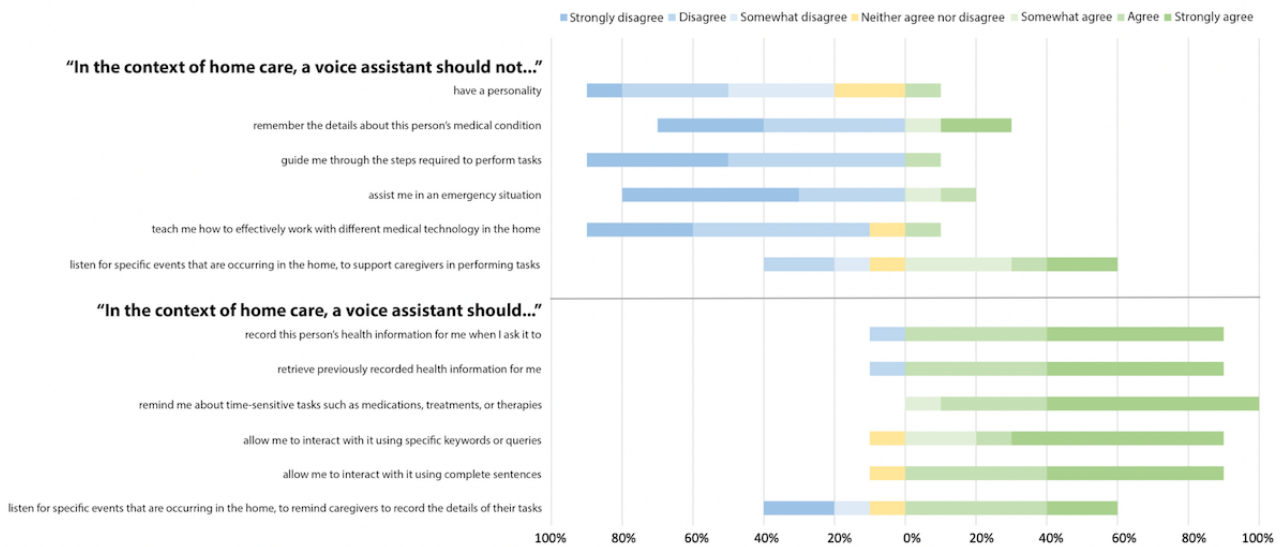
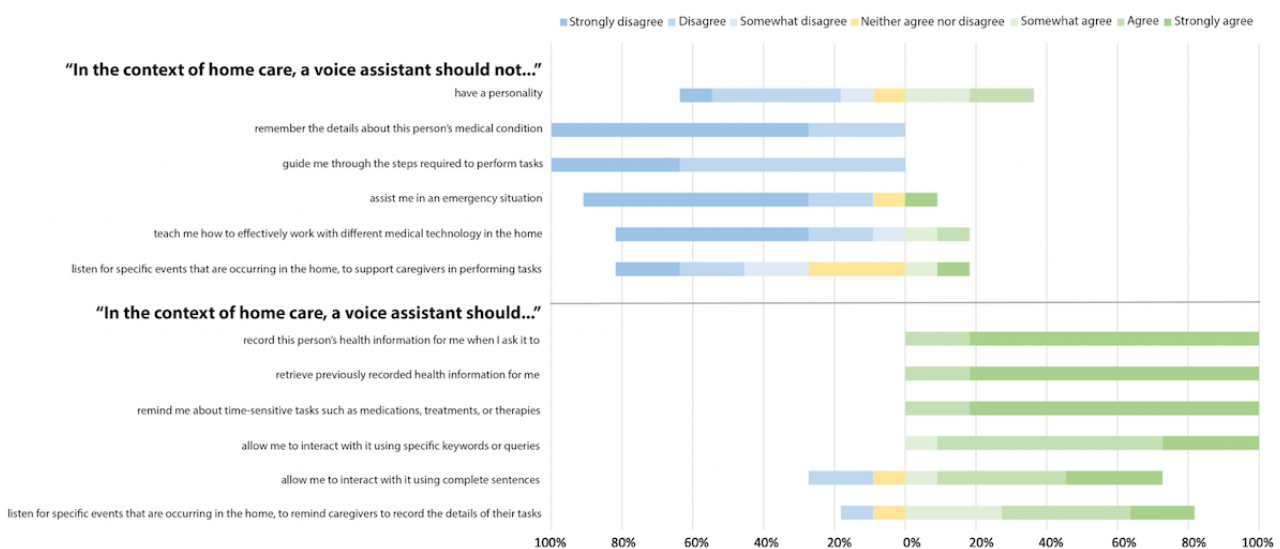


Figure 4. Caregiver expectations of voice assistants in complex home care—minimal experience (n=11).



VAs With a Personality

VA personality may be attributed to cognitive, emotional, and social human characteristics [45]. However, the participants in this study were not provided with an official definition when responding to the personality-based Likert-scale question, which

may have influenced the degree of contrasting perspectives that was observed. Regardless, the participants often qualified their responses, providing clarification on their expectations of this quantitative measure.

The participants who agreed that a VA should have a personality expressed that they would want it to be happy and positive. A

participant qualified that VAs provide objective responses; therefore, it would be acceptable for a VA to express a personality. Although most participants agreed that VAs should have a personality, more participants who had minimal experience with VAs agreed that a VA should not have a personality (4/11, 36%) compared with participants who had expressed having more experience using VAs (1/10, 10%). The remaining participants (3/22, 14%) were unsure about this potential feature.

The participants who disagreed about VAs having a personality were particularly concerned about the influence that a personality from a device could have on vulnerable caregivers in specific contexts. For example, a *happy* demeanor in a VA providing information about missing scheduled medications could inappropriately influence caregivers' attitudes toward medication adherence. A participant argued that the vulnerability of caregivers should therefore play a role in designing a VA personality for home care:

I strongly feel that it shouldn't have a personality...I think that could take advantage of vulnerable people...I know that there's an argument to be made the exact opposite—that it would make it more user-friendly, it would make it warmer, it could be a companion to the person, etc. There's a lot of lines you can cross... [Participant 7, minimal experience]

They observed the benefit of a VA personality being potentially more user-friendly and acting as a caregiver itself. However, the concept of a VA displaying emotion may inadvertently manipulate the caregivers' perception of care, which could be particularly harmful in medically fragile situations.

Assisting in an Emergency

Most participants (18/22, 82%) agreed that VAs should assist caregivers in an emergency, and this outcome was relatively balanced between experienced and minimally experienced participants (Figures 3 and 4). However, the participants who disagreed about this potential VA feature expressed that the ability of a VA to assist in an emergency should be limited to calling emergency services (eg, calling 911). Calling for the help of another human should be the extent of a VA's support in this type of situation:

[With] the medical conditions my [child] has...I don't think I'm anywhere near trusting a device...Yeah, not yet. [Participant 9, experienced]

A home care situation may be too complicated for a VA to provide help if there is an emergency. There are likely several factors of the environment and the situation with the child or older adult that the VA cannot perceive. Trust was a concern for some participants (2/22, 9%) in this context, where the VA would need to be 100% accurate in its response if they were to trust it completely.

Teaching and Guiding Caregivers Through Tasks

Guiding caregivers through tasks in the home was a potential VA functionality that many caregivers (21/22, 95%) mentioned should be supported. However, other participants (4/22, 18%)

also noted that VAs should not be *initially* teaching caregivers how to perform tasks that they have never done before:

Some of my hesitation was that I was defaulting to the importance of face-to-face. If you're training a new nurse, from my experience, you want someone there on the premises training you in-person: One, for the registered staff to have confidence in the new person, new trainees' ability, but also, I would think to instill more confidence in the patient in the new caregiver. [Participant 5, minimal experience]

The participants emphasized the importance of having in-person training and the need to set access limitations for specific caregiver populations, especially in learning how to use a device that dispenses medication. There was also a concern about a VA providing information about accessing medication-dispensing equipment that could endanger patient safety.

Listening to Events in the Home

An always-on VA capable of unprompted responses was seen either as a privacy issue or as significant support for home care safety. Concerning privacy, the participants expressed that they did not like the idea of VAs being present and having the ability to speak without previous notice. Although the participants mentioned that they observed the VA's potential to notify them about safety events concerning the care situation, other participants said that they would not be comfortable with unprompted interactions. If the VA could respond without being prompted, the participants expressed that this would be an invasion of the private activities in their homes:

In some situations, that could be of significant support and...some situations, that might also be like an invasion of privacy. [Participant 2, minimal experience]

The participants expressed that unprompted responses from the VAs would support peace of mind for their respite care concerning safety. A participant described that unprompted responses from a VA could be used to remind their PSW where to stand when performing physical therapy with their spouse:

That would be great for me because I'm not in the room when these caregivers come, and they're going to be the ones to tell them to stand behind [my spouse]. [Participant 17, minimal experience]

Finally, a participant mentioned that VAs could listen for unexpected accidents in the home, such as a fall, and promptly notify caregivers to act on issues. They also noted the potential for VAs to identify caregiver abuse:

That could be a huge safety component...to identify caregiver abuse...because really there is caregiver abuse... [Participant 13, minimal experience]

Discussion

Principal Findings

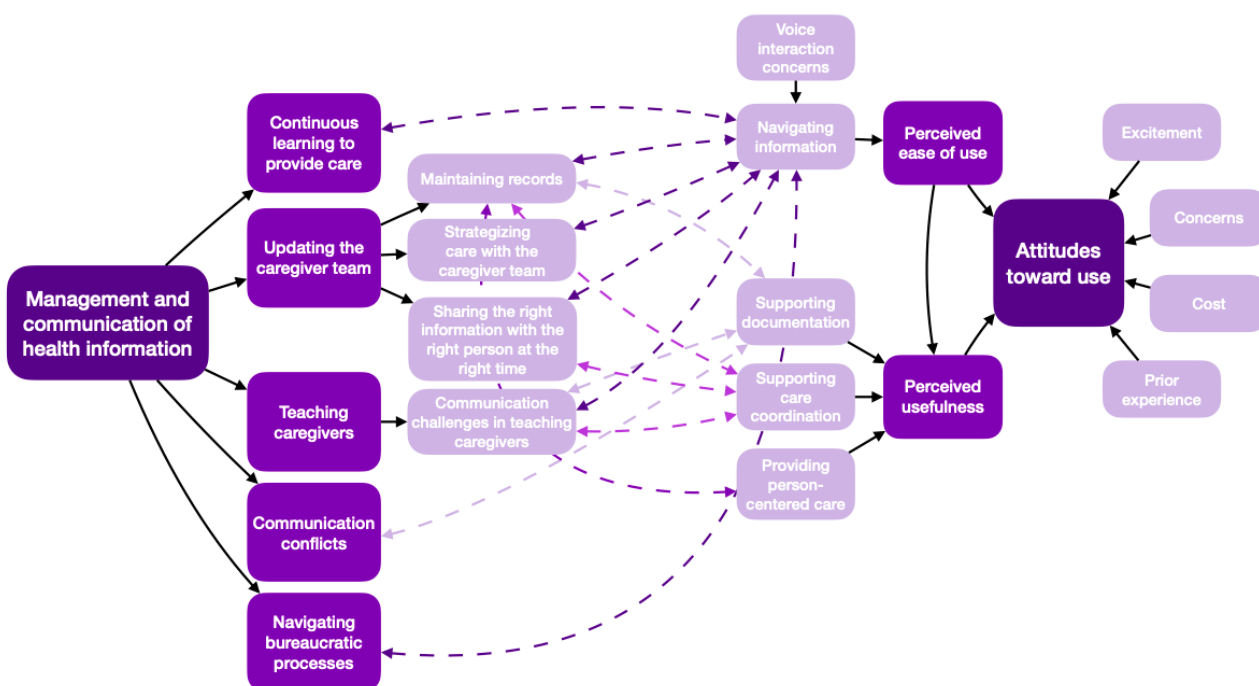
This study captured the initial perspectives of a sample of caregivers regarding the acceptance of VAs to inform digital technology design for complex home care. This study identified

the importance of utility and ease of interaction in influencing technology adoption. The expectations for VAs to support caregivers in managing and communicating health information may positively affect caregivers' desire to integrate VAs into complex home care while being influenced by previous experiences using VAs. Triangulation of qualifying responses with the Likert-scale results also identified critical design concerns and ethical considerations for using VAs to support caregiving. In the following sections, we discuss the importance of designing VAs for usefulness, ease of use, and the context within which a VA may be used in complex home care.

Designing for Usefulness

Previous research on complex home care has formatively identified some of the high-level health information management and communication processes of caregivers in the context of children with special health care needs [26]. With these previous findings and the outcomes of this study on caregivers' initial beliefs about the design functionalities that VAs could provide in complex home care, we can begin to map the design of VA technology to the home care work domain. Several connections can be made between caregiving tasks and caregivers' perspectives regarding the ease of use and usefulness of VAs for complex home care that may ultimately influence their attitudes toward integrating this technology (Figure 5).

Figure 5. Mapping information management and communication processes with expectations of voice assistants in complex home care.



The utility of VAs is one of the primary motivators for individuals to use this technology in their home [46]. However, understanding the specific factors influencing their utility requires a deeper understanding to inform practical guidelines for developers [46]. Our study can begin to inform the factors influencing VA utility for its use in complex home care. Caregiver participants perceive the utility of VAs in the context of documentation, care coordination, and provision of person-centered care.

Similar to the findings of Sezgin et al [7], this study identified that both family caregiver and hired caregiver participants felt positively toward VAs in the context of recording health information by voice. Family caregiver participants particularly expected VAs to improve the organization of records that could be quickly updated and retrieved in their homes. The utility provided through organized, accessible information could reduce the burden on caregivers to communicate information to others. It may also alleviate conflicts from miscommunication with other caregivers when caregiving teams are large. Furthermore, rather than relying on the primary caregiver to ensure that

documented updates are communicated to incoming caregivers, the VA could be used to communicate health information summaries, trends, and other necessary documentation to caregivers. This functionality could relieve the primary caregiver from having to report these details repeatedly.

Ensuring that complete and accurate information is shared about home care can be a challenging task when there are multiple caregivers involved with various responsibilities [25]. An important task that VAs could support is shift handoffs for hired caregivers. A VA for home care built into a mobile app or smart speaker could capitalize on location and scheduling information to provide context-specific details about previous care activities that occurred in the home, supporting a caregiver's situational awareness before the beginning of their shift. Verbal reminders from a VA located in the house about time-sensitive tasks could further support caregivers' memory to perform specific tasks or track information when providing care to multiple individuals in a single day.

The participants in this study particularly observed the utility of VAs interacting with the care receiver. Although this was not the focus of this study, the participants identified the impact that VAs could have on supporting self-management of care in the home for children with special health care needs and older adults. The proposed use of VAs by these populations has been previously identified by Sezgin et al [12,13] for applications such as medication tracking under parental, guardian, or caregiver supervision and other health documentation tasks [7]. Research has been directly conducted on children's interactions with VAs, providing evidence for their positive interactions with this technology in general contexts [47-50]. The participants in this study explained that a VA could provide emotional support to children with special health care needs during potentially uncomfortable medical procedures or for medication adherence, supporting caregiving beyond aspects of health information management and communication. Future research should explore the use of VAs to help children and older adults in these contexts.

Designing for Ease of Use

In complex home care, the exchange of information is fundamental to the successful outcomes of learning, sharing knowledge, teaching, communicating, and the bureaucratic processing of caregiving [51-53]. Information is dynamically flowing among caregivers within these subdomains of the work environment. However, the information structure in paper-based systems or mobile apps can make accessing it challenging [26]. One of the primary benefits that VAs provide compared with graphical or other physical user interfaces is the removal of visual hierarchies by accessing information through voice commands [13,34].

However, the findings from this study suggest that there is an influencing factor of ease-of-use concerns that may negatively affect information navigation through voice controls. If the VA cannot understand a user's speech, the ease of use will be severely affected in relation to being error prone. VA technology that supports individuals with speech impairments, such as Google's *Project Euphonia*, is a critical research area that should be appropriately addressed to successfully integrate this technology into complex home care [54].

Designing for Use Context: VA Personality

With their inherent communication mode being conversational and potentially human-like—attributing it to cognitive, emotional, and social human characteristics [45]—it is crucial to consider the potential influence of VA personality on the use of this technology [46]. Baptista et al [55] previously identified that personality could influence the users' perceptions of a VA's role in health care. In their study, participants perceived the personality of an embodied VA for diabetes management as a friendly coach more than a health professional [55]. A scoping review by Car et al [56] identified other personality traits in studies with VAs in a health care context: informal, human-like, culture-specific, factual, gender-specific, and conversational agent. Given the exploratory nature of our study, the participants were not provided with an initial definition of VA personality or examples of what the personality of a home care VA could

be when answering this Likert-scale question, which may have influenced the differences in their responses.

As a result of this nongrounded approach, the use context was identified as a critical factor in caregiver participants' expectations of a home care VA expressing potential cognitive, emotional, or social characteristics. When designing VA personalities for family caregivers, this population can be considered vulnerable; it is essential to consider the influence of personality traits on their reliance on this type of technology in different caregiving situations. Although the design of VAs currently includes human-like personality traits for health care applications in specific contexts, such as adherence to active living regimens and psychological difficulties [57,58], the participants in this study expected VAs to assist in more than one context. A consistent personality trait for VAs may not be appropriate for every home care situation and may negatively influence a caregiver's perception. Future research should explore how personality traits influence caregiver engagement, reliance on technology, and medical decision-making.

Designing for Use Context: Intelligent Support

The extent to which the participants in this study initially expected a VA to assist their tasks suggests that caregivers might prefer a less intelligent VA that is limited to providing a means for retrieving previously entered information. Insights into or interpretations of health information may be an unexpected output from a VA by caregivers while also posing a risk of adverse events [13]. The caregiver participants in this study mainly wished to direct the interaction with VA technology, where the information exchange was not expected to advance beyond their initial intents. Intelligent VAs may be better integrated as complementary caregiving tools [59]. For example, our participants discussed using a VA to create reminders or instructions for procedures based on the information they would consciously provide to the system. When they need assistance, they would prefer to contact other caregivers through the VA rather than asking the VA itself to assist them despite its potential knowledge base.

Finally, although context-specific interactions may improve engagement and adoption of VAs by general consumers, this functionality may require predictive algorithms based on enormous amounts of data about the home to support the system's intelligence [60,61]. With the uniqueness of the participants' caregiving backgrounds and home care experiences, some participants would be positively inclined toward a VA that provides context-specific support through passive information collection. However, collecting audio data about the home environment raises ethical considerations. It is essential to consider how these data are used to report home events ranging from accidents to potential caregiver abuse, especially for user groups who find it challenging to navigate the complexities of security choices for Internet-of-Things devices [62]. Caregivers may be concerned if there is no option to control the always listening and analyzing functionality [63].

Strengths and Limitations

The nature of this exploratory study on the participants' initial expectations of using VAs in complex home care captured the

unique perspectives of the potential primary users of this technology. There has been no significant research conducted using the TAM for understanding VA acceptance and none captured during COVID-19. Although demographics are limited, they offer preliminary insights into diverse situations.

Future work should expand on these results to examine more viewpoints, including people being cared for, various health care professionals, regulators, and technology experts, ultimately bringing a holistic understanding of the system itself and its potential. Additional studies should also examine the potential of VA personality with respect to specific cognitive, emotional, and social human-like attributes and its impact on caregiver perceptions of care, as well as the potential of other methods of conversational interaction with digital tools such as text-based or visual interfaces. An increased sample size through further research would provide more insight into differing caregivers' perspectives on VAs in complex home care.

Conclusions

This study provides early emerging research into understanding caregiver perspectives on VAs to support complex home care

using the TAM supplemented by a Likert-scale questionnaire. The results point toward the factors influencing the utility of VAs in this work domain and how the ease of interacting with health information through a VA may influence technology adoption. VAs could provide utility for caregivers' current health care documentation methods and care coordination in the home. There is a desire for VAs to support care recipient independence in the contexts of children with special health care needs and older adults beyond the aspects of information management, providing opportunities for further studies.

Beyond health information interaction, there are ethical considerations for using a VA that provides contextually specific insights from collected audio data given the complexity and diversity of activities occurring in the home. The design of a VA personality should carefully evaluate its potential influence on vulnerable caregiver populations' perceptions of care. Future research should focus on integrating VAs into specific contexts of information management and communication for complex home care to further understand the factors influencing utility, ease of use, and adoption in the design of this technology.

Acknowledgments

The authors are grateful for the generosity of their participants with their time and insights. This study was funded by Telus Health and a Natural Sciences and Engineering Research Council of Canada Collaborative Research and Development grant, as well as the Ontario Graduate Scholarship and the University of Waterloo Alumni@Microsoft Graduate Scholarship.

Conflicts of Interest

None declared.

References

1. Turing AM. I.—Computing machinery and intelligence. *Mind* 1950 Oct;LIX(236):433-460. [doi: [10.1093/mind/LIX.236.433](https://doi.org/10.1093/mind/LIX.236.433)]
2. Roe DB, Wilpon JG. *Voice Communication Between Humans and Machines*. Washington, DC, USA: National Academies Press (US); 1994:31.
3. Clark L, Doyle P, Garaialde D, Gilmartin E, Schlögl S, Edlund J, et al. The state of speech in HCI: trends, themes and challenges. *Interact Comput* 2019 Jun;31(4):349-371. [doi: [10.1093/iwc/iwz016](https://doi.org/10.1093/iwc/iwz016)]
4. Hoy MB. Alexa, Siri, Cortana, and more: an introduction to voice assistants. *Med Ref Serv Q* 2018;37(1):81-88. [doi: [10.1080/02763869.2018.1404391](https://doi.org/10.1080/02763869.2018.1404391)] [Medline: [29327988](https://pubmed.ncbi.nlm.nih.gov/29327988/)]
5. Benlian A, Klumpe J, Hinz O. Mitigating the intrusive effects of smart home assistants by using anthropomorphic design features: a multimethod investigation. *Inf Syst J* 2020 Nov;30(6):1010-1042. [doi: [10.1111/isj.12243](https://doi.org/10.1111/isj.12243)]
6. Pitardi V, Marriott HR. Alexa, she's not human but... Unveiling the drivers of consumers' trust in voice-based artificial intelligence. *Psychol Mark* 2021 Jan 20;38(4):626-642. [doi: [10.1002/mar.21457](https://doi.org/10.1002/mar.21457)]
7. Sezgin E, Noritz G, Lin S, Huang Y. Feasibility of a voice-enabled medical diary app (SpeakHealth) for caregivers of children with special health care needs and health care providers: mixed methods study. *JMIR Form Res* 2021 May 11;5(5):e25503 [FREE Full text] [doi: [10.2196/25503](https://doi.org/10.2196/25503)] [Medline: [33865233](https://pubmed.ncbi.nlm.nih.gov/33865233/)]
8. Li J, Maharjan B, Xie B, Tao C. A personalized voice-based diet assistant for caregivers of Alzheimer disease and related dementias: system development and validation. *J Med Internet Res* 2020 Sep 21;22(9):e19897 [FREE Full text] [doi: [10.2196/19897](https://doi.org/10.2196/19897)] [Medline: [32955452](https://pubmed.ncbi.nlm.nih.gov/32955452/)]
9. Jesús-Azabal M, Rojo J, Moguel E, Flores-Martin D, Berrocal J, García-Alonso J, et al. Voice assistant to remind pharmacologic treatment in elders. In: *Proceedings of the 2nd International Workshop on Gerontechnology*. 2019 Presented at: IWoG '19; September 4–5, 2019; Cáceres, Spain p. 123-133. [doi: [10.1007/978-3-030-41494-8_12](https://doi.org/10.1007/978-3-030-41494-8_12)]
10. Sezgin E, Huang Y, Ramtekkar U, Lin S. Readiness for voice assistants to support healthcare delivery during a health crisis and pandemic. *NPJ Digit Med* 2020 Sep 16;3:122 [FREE Full text] [doi: [10.1038/s41746-020-00332-0](https://doi.org/10.1038/s41746-020-00332-0)] [Medline: [33015374](https://pubmed.ncbi.nlm.nih.gov/33015374/)]
11. Nasr M, Karray F, Quintana Y. Human machine interaction platform for home care support system. In: *Proceedings of the 2020 IEEE International Conference on Systems, Man, and Cybernetics*. 2020 Presented at: SMC '20; October 11-14, 2020; Toronto, Canada p. 4210-4215. [doi: [10.1109/smc42975.2020.9283095](https://doi.org/10.1109/smc42975.2020.9283095)]

12. Sezgin E, Weiler M, Weiler A, Lin S. Proposing an ecosystem of digital health solutions for teens with chronic conditions transitioning to self-management and independence: exploratory qualitative study. *J Med Internet Res* 2018 Sep 06;20(9):e10285 [FREE Full text] [doi: [10.2196/10285](https://doi.org/10.2196/10285)] [Medline: [30190253](https://pubmed.ncbi.nlm.nih.gov/30190253/)]
13. Sezgin E, Noritz G, Elek A, Conkol K, Rust S, Bailey M, et al. Capturing at-home health and care information for children with medical complexity using voice interactive technologies: multi-stakeholder viewpoint. *J Med Internet Res* 2020 Feb 13;22(2):e14202 [FREE Full text] [doi: [10.2196/14202](https://doi.org/10.2196/14202)] [Medline: [32053114](https://pubmed.ncbi.nlm.nih.gov/32053114/)]
14. Lindt N, van Berkel J, Mulder BC. Determinants of overburdening among informal carers: a systematic review. *BMC Geriatr* 2020 Aug 26;20(1):304 [FREE Full text] [doi: [10.1186/s12877-020-01708-3](https://doi.org/10.1186/s12877-020-01708-3)] [Medline: [32847493](https://pubmed.ncbi.nlm.nih.gov/32847493/)]
15. Szlenk-Czyczerska E, Guzek M, Bielska DE, Ławnik A, Polański P, Kurpas D. Needs, aggravation, and degree of burnout in informal caregivers of patients with chronic cardiovascular disease. *Int J Environ Res Public Health* 2020 Sep 03;17(17):6427 [FREE Full text] [doi: [10.3390/ijerph17176427](https://doi.org/10.3390/ijerph17176427)] [Medline: [32899352](https://pubmed.ncbi.nlm.nih.gov/32899352/)]
16. Ennis A, Rafferty J, Synnott J, Cleland I, Nugent C, Selby A, et al. A smart cabinet and voice assistant to support independence in older adults. In: Proceedings of the 11th International Conference on the Ubiquitous Computing and Ambient Intelligence. 2017 Presented at: UCAmI '17; November 7–10, 2017; Philadelphia, PA, USA p. 466-472. [doi: [10.1007/978-3-319-67585-5_47](https://doi.org/10.1007/978-3-319-67585-5_47)]
17. Sanders J, Martin-Hammond A. Exploring autonomy in the design of an intelligent health assistant for older adults. In: Proceedings of the 24th International Conference on Intelligent User Interfaces: Companion. 2019 Presented at: IUI '19; March 16-20, 2019; Marina del Ray, CA, USA p. 95-96. [doi: [10.1145/3308557.3308713](https://doi.org/10.1145/3308557.3308713)]
18. Choi YK, Thompson HJ, Demiris G. Use of an Internet-of-Things smart home system for healthy aging in older adults in residential settings: pilot feasibility study. *JMIR Aging* 2020 Nov 10;3(2):e21964 [FREE Full text] [doi: [10.2196/21964](https://doi.org/10.2196/21964)] [Medline: [33170128](https://pubmed.ncbi.nlm.nih.gov/33170128/)]
19. Chin J, Quinn K, Muramatsu N, Marquez D. A user study on the feasibility and acceptance of delivering physical activity programs to older adults through conversational agents. *Proc Hum Factors Ergon Soc Annu Meet* 2020 Dec;64(1):33-37 [FREE Full text] [doi: [10.1177/1071181320641010](https://doi.org/10.1177/1071181320641010)] [Medline: [34177216](https://pubmed.ncbi.nlm.nih.gov/34177216/)]
20. Caregiving in the U.S. 2020: Infographic. National Alliance for Caregiving and American Association of Retired Persons. 2020 May. URL: <https://www.caregiving.org/wp-content/uploads/2020/05/AARP-835-AARP-Caregiving-in-the-US-Infographics-vFINAL-1.pdf> [accessed 2022-02-02]
21. Verloo H, Kempel T, Vidal N, Pereira F. Perceptions about technologies that help community-dwelling older adults remain at home: qualitative study. *J Med Internet Res* 2020 Jun 04;22(6):e17930 [FREE Full text] [doi: [10.2196/17930](https://doi.org/10.2196/17930)] [Medline: [32496197](https://pubmed.ncbi.nlm.nih.gov/32496197/)]
22. Lang A, Macdonald M, Marck P, Toon L, Griffin M, Easty T, et al. Seniors managing multiple medications: using mixed methods to view the home care safety lens. *BMC Health Serv Res* 2015 Dec 12;15:548 [FREE Full text] [doi: [10.1186/s12913-015-1193-5](https://doi.org/10.1186/s12913-015-1193-5)] [Medline: [26651331](https://pubmed.ncbi.nlm.nih.gov/26651331/)]
23. Desai AD, Durkin LK, Jacob-Files EA, Mangione-Smith R. Caregiver perceptions of hospital to home transitions according to medical complexity: a qualitative study. *Acad Pediatr* 2016 Mar;16(2):136-144. [doi: [10.1016/j.acap.2015.08.003](https://doi.org/10.1016/j.acap.2015.08.003)] [Medline: [26703883](https://pubmed.ncbi.nlm.nih.gov/26703883/)]
24. Nasarwanji M, Werner NE, Carl K, Hohl D, Leff B, Gurses AP, et al. Identifying challenges associated with the care transition workflow from hospital to skilled home health care: perspectives of home health care agency providers. *Home Health Care Serv Q* 2015;34(3-4):185-203. [doi: [10.1080/01621424.2015.1092908](https://doi.org/10.1080/01621424.2015.1092908)] [Medline: [26495858](https://pubmed.ncbi.nlm.nih.gov/26495858/)]
25. Tang C, Chen Y, Cheng K, Ngo V, Mattison JE. Awareness and handoffs in home care: coordination among informal caregivers. *Behav Inf Technol* 2018;37(1):66-86. [doi: [10.1080/0144929x.2017.1405073](https://doi.org/10.1080/0144929x.2017.1405073)]
26. Tennant R, Allana S, Mercer K, Burns CM. Exploring the experiences of family caregivers of children with special health care needs to inform the design of digital health systems: formative qualitative study. *JMIR Form Res* 2022 Jan 05;6(1):e28895 [FREE Full text] [doi: [10.2196/28895](https://doi.org/10.2196/28895)] [Medline: [34989692](https://pubmed.ncbi.nlm.nih.gov/34989692/)]
27. Guisado-Fernandez E, Caulfield B, Silva PA, Mackey L, Singleton D, Leahy D, et al. Development of a caregivers' support platform (connected health sustaining home stay in dementia): protocol for a longitudinal observational mixed methods study. *JMIR Res Protoc* 2019 Aug 28;8(8):13280 [FREE Full text] [doi: [10.2196/13280](https://doi.org/10.2196/13280)] [Medline: [31464187](https://pubmed.ncbi.nlm.nih.gov/31464187/)]
28. Godwin KM, Mills WL, Anderson JA, Kunik ME. Technology-driven interventions for caregivers of persons with dementia: a systematic review. *Am J Alzheimers Dis Other Demen* 2013 May;28(3):216-222 [FREE Full text] [doi: [10.1177/1533317513481091](https://doi.org/10.1177/1533317513481091)] [Medline: [23528881](https://pubmed.ncbi.nlm.nih.gov/23528881/)]
29. Yamashita N, Kuzuoka H, Kudo T, Hirata K, Aramaki E, Hattori K. How information sharing about care recipients by family caregivers impacts family communication. In: Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems. 2018 Presented at: CHI '18; April 21-26, 2018; Montreal, Canada p. 1-13. [doi: [10.1145/3173574.3173796](https://doi.org/10.1145/3173574.3173796)]
30. Wang J, Howell D, Shen N, Geng Z, Wu F, Shen M, et al. Mhealth supportive care intervention for parents of children with acute lymphoblastic leukemia: quasi-experimental pre- and postdesign study. *JMIR Mhealth Uhealth* 2018 Nov 19;6(11):e195 [FREE Full text] [doi: [10.2196/mhealth.9981](https://doi.org/10.2196/mhealth.9981)] [Medline: [30455166](https://pubmed.ncbi.nlm.nih.gov/30455166/)]

31. Heynsbergh N, Heckel L, Botti M, Livingston PM. A smartphone app to support carers of people living with cancer: a feasibility and usability study. *JMIR Cancer* 2019 Jan 31;5(1):e11779 [FREE Full text] [doi: [10.2196/11779](https://doi.org/10.2196/11779)] [Medline: [30702432](https://pubmed.ncbi.nlm.nih.gov/30702432/)]
32. Kumah-Crystal YA, Pirtle CJ, Whyte HM, Goode ES, Anders SH, Lehmann CU. Electronic health record interactions through voice: a review. *Appl Clin Inform* 2018 Jul;9(3):541-552 [FREE Full text] [doi: [10.1055/s-0038-1666844](https://doi.org/10.1055/s-0038-1666844)] [Medline: [30040113](https://pubmed.ncbi.nlm.nih.gov/30040113/)]
33. Kim S. Exploring how older adults use a smart speaker-based voice assistant in their first interactions: qualitative study. *JMIR Mhealth Uhealth* 2021 Jan 13;9(1):e20427 [FREE Full text] [doi: [10.2196/20427](https://doi.org/10.2196/20427)] [Medline: [33439130](https://pubmed.ncbi.nlm.nih.gov/33439130/)]
34. Nallam P, Bhandari S, Sanders J, Martin-Hammond A. A question of access: exploring the perceived benefits and barriers of intelligent voice assistants for improving access to consumer health resources among low-income older adults. *Gerontol Geriatr Med* 2020 Dec 29;6:2333721420985975 [FREE Full text] [doi: [10.1177/2333721420985975](https://doi.org/10.1177/2333721420985975)] [Medline: [33457459](https://pubmed.ncbi.nlm.nih.gov/33457459/)]
35. Corbett CF, Combs EM, Wright PJ, Owens OL, Stringfellow I, Nguyen T, et al. Virtual home assistant use and perceptions of usefulness by older adults and support person dyads. *Int J Environ Res Public Health* 2021 Jan 27;18(3):1113 [FREE Full text] [doi: [10.3390/ijerph18031113](https://doi.org/10.3390/ijerph18031113)] [Medline: [33513798](https://pubmed.ncbi.nlm.nih.gov/33513798/)]
36. Marikyan D, Papagiannidis S, Alamanos E. A systematic review of the smart home literature: a user perspective. *Technol Forecast Soc Change* 2019 Jan;138:139-154. [doi: [10.1016/j.techfore.2018.08.015](https://doi.org/10.1016/j.techfore.2018.08.015)]
37. Hong YY, Morris MW, Chiu CY, Benet-Martínez V. Multicultural minds. A dynamic constructivist approach to culture and cognition. *Am Psychol* 2000 Jul;55(7):709-720. [doi: [10.1037//0003-066x.55.7.709](https://doi.org/10.1037//0003-066x.55.7.709)] [Medline: [10916861](https://pubmed.ncbi.nlm.nih.gov/10916861/)]
38. Oliver C, Nesbit S, Kelly N. Dissolving dualisms: how two positivists engaged with non-positivist qualitative methodology. *Int J Qual Methods* 2013 Feb 01;12(1):180-194. [doi: [10.1177/160940691301200106](https://doi.org/10.1177/160940691301200106)]
39. Gale NK, Heath G, Cameron E, Rashid S, Redwood S. Using the framework method for the analysis of qualitative data in multi-disciplinary health research. *BMC Med Res Methodol* 2013 Sep 18;13:117 [FREE Full text] [doi: [10.1186/1471-2288-13-117](https://doi.org/10.1186/1471-2288-13-117)] [Medline: [24047204](https://pubmed.ncbi.nlm.nih.gov/24047204/)]
40. Braun V, Clarke V. Using thematic analysis in psychology. *Qual Res Psychol* 2006 Jan;3(2):77-101. [doi: [10.1191/1478088706qp063oa](https://doi.org/10.1191/1478088706qp063oa)]
41. Davis FD. Perceived usefulness, perceived ease of use, and user acceptance of information technology. *MIS Q* 1989 Sep;13(3):319-340. [doi: [10.2307/249008](https://doi.org/10.2307/249008)]
42. Rahimi B, Nadri H, Lotfnezhad Afshar H, Timpka T. A systematic review of the technology acceptance model in health informatics. *Appl Clin Inform* 2018 Jul;9(3):604-634 [FREE Full text] [doi: [10.1055/s-0038-1668091](https://doi.org/10.1055/s-0038-1668091)] [Medline: [30112741](https://pubmed.ncbi.nlm.nih.gov/30112741/)]
43. Nadri H, Rahimi B, Lotfnezhad Afshar H, Samadbeik M, Garavand A. Factors affecting acceptance of hospital information systems based on extended technology acceptance model: a case study in three paraclinical departments. *Appl Clin Inform* 2018 Apr;9(2):238-247 [FREE Full text] [doi: [10.1055/s-0038-1641595](https://doi.org/10.1055/s-0038-1641595)] [Medline: [29618139](https://pubmed.ncbi.nlm.nih.gov/29618139/)]
44. Nielsen J. What is usability? In: Nielsen J, editor. *Usability Engineering*. Burlington, MA, USA: Morgan Kaufman; 1993:23-48.
45. Poushneh A. Humanizing voice assistant: the impact of voice assistant personality on consumers' attitudes and behaviors. *J Retail Consum Serv* 2021 Jan;58:102283. [doi: [10.1016/j.jretconser.2020.102283](https://doi.org/10.1016/j.jretconser.2020.102283)]
46. McLean G, Osei-Frimpong K. Hey Alexa ... examine the variables influencing the use of artificial intelligent in-home voice assistants. *Comput Human Behav* 2019 Oct;99:28-37. [doi: [10.1016/j.chb.2019.05.009](https://doi.org/10.1016/j.chb.2019.05.009)]
47. Festerling J, Siraj I. Alexa, what are you? Exploring primary school children's ontological perceptions of digital voice assistants in open interactions. *Human Dev* 2020 Jun 5;64(1):26-43. [doi: [10.1159/000508499](https://doi.org/10.1159/000508499)]
48. Ligthart M, Van Bindsbergen KL, Fernhout T, Grootenhuis MA, Neerincx MA, Hindriks KV. A child and a robot getting acquainted - interaction design for eliciting self-disclosure. In: *Proceedings of the 18th International Conference on Autonomous Agents and Multiagent Systems*. 2019 Presented at: AAMAS '19; May 13-17, 2019; Montreal, Canada p. 61-70.
49. Cheng Y, Yen K, Chen Y, Chen S, Hiniker A. Why doesn't it work?: voice-driven interfaces and young children's communication repair strategies. In: *Proceedings of the 17th ACM Conference on Interaction Design and Children*. 2018 Jun Presented at: IDC '18; June 19-22, 2018; Trondheim, Norway p. 337-348. [doi: [10.1145/3202185.3202749](https://doi.org/10.1145/3202185.3202749)]
50. Aeschlimann S, Bleiker M, Wechner M, Gampe A. Communicative and social consequences of interactions with voice assistants. *Comput Human Behav* 2020 Nov;112:106466. [doi: [10.1016/j.chb.2020.106466](https://doi.org/10.1016/j.chb.2020.106466)]
51. Foged S, Nørholm V, Andersen O, Petersen HV. Nurses' perspectives on how an e-message system supports cross-sectoral communication in relation to medication administration: a qualitative study. *J Clin Nurs* 2018 Feb;27(3-4):795-806. [doi: [10.1111/jocn.14109](https://doi.org/10.1111/jocn.14109)] [Medline: [29048777](https://pubmed.ncbi.nlm.nih.gov/29048777/)]
52. Koru G, Alhuwail D, Topaz M, Norcio AF, Mills ME. Investigating the challenges and opportunities in home care to facilitate effective information technology adoption. *J Am Med Dir Assoc* 2016 Jan;17(1):53-58. [doi: [10.1016/j.jamda.2015.10.008](https://doi.org/10.1016/j.jamda.2015.10.008)] [Medline: [26612483](https://pubmed.ncbi.nlm.nih.gov/26612483/)]
53. Alvarado CS, Zook K, Henry J. *ONC Data Brief No. 39 – Electronic Health Record Adoption and Interoperability among U.S. Skilled Nursing Facilities in 2016*. Office of the National Coordinator for Health Information Technology. 2017 Sep. URL: <https://www.healthit.gov/sites/default/files/electronic-health-record-adoption-and-interoperability-among-u.s.-skilled-nursing-facilities-in-2016.pdf> [accessed 2022-02-02]

54. Ostherr K. Artificial intelligence and medical humanities. *J Med Humanit* (forthcoming) 2020 Jul 11. [doi: [10.1007/s10912-020-09636-4](https://doi.org/10.1007/s10912-020-09636-4)] [Medline: [32654043](https://pubmed.ncbi.nlm.nih.gov/32654043/)]
55. Baptista S, Wadley G, Bird D, Oldenburg B, Speight J, My Diabetes Coach Research Group. Acceptability of an embodied conversational agent for type 2 diabetes self-management education and support via a smartphone app: mixed methods study. *JMIR Mhealth Uhealth* 2020 Jul 22;8(7):e17038 [FREE Full text] [doi: [10.2196/17038](https://doi.org/10.2196/17038)] [Medline: [32706734](https://pubmed.ncbi.nlm.nih.gov/32706734/)]
56. Tudor Car L, Dhinakaran DA, Kyaw BM, Kowatsch T, Joty S, Theng Y, et al. Conversational agents in health care: scoping review and conceptual analysis. *J Med Internet Res* 2020 Aug 07;22(8):e17158 [FREE Full text] [doi: [10.2196/17158](https://doi.org/10.2196/17158)] [Medline: [32763886](https://pubmed.ncbi.nlm.nih.gov/32763886/)]
57. Davis CR, Murphy KJ, Curtis RG, Maher CA. A process evaluation examining the performance, adherence, and acceptability of a physical activity and diet artificial intelligence virtual health assistant. *Int J Environ Res Public Health* 2020 Dec 07;17(23):9137 [FREE Full text] [doi: [10.3390/ijerph17239137](https://doi.org/10.3390/ijerph17239137)] [Medline: [33297456](https://pubmed.ncbi.nlm.nih.gov/33297456/)]
58. Falala-Séchet C, Antoine L, Thiriez I, Bungener C. Owlle: a chatbot that provides emotional support for coping with psychological difficulties. In: *Proceedings of the 19th ACM International Conference on Intelligent Virtual Agents*. 2019 Presented at: IVA '19; July 2-5, 2019; Paris, France p. 236-237. [doi: [10.1145/3308532.3329416](https://doi.org/10.1145/3308532.3329416)]
59. Powell J. Trust me, I'm a chatbot: how artificial intelligence in health care fails the Turing test. *J Med Internet Res* 2019 Oct 28;21(10):e16222 [FREE Full text] [doi: [10.2196/16222](https://doi.org/10.2196/16222)] [Medline: [31661083](https://pubmed.ncbi.nlm.nih.gov/31661083/)]
60. Pal D, Arpnikanondt C, Funilkul S, Chutimaskul W. The adoption analysis of voice-based smart IoT products. *IEEE Internet Things J* 2020 May 1;7(11):10852-10867. [doi: [10.1109/jiot.2020.2991791](https://doi.org/10.1109/jiot.2020.2991791)]
61. Moriuchi E. Okay, Google!: an empirical study on voice assistants on consumer engagement and loyalty. *Psychol Mark* 2019 Jan 15;36(5):489-501. [doi: [10.1002/mar.21192](https://doi.org/10.1002/mar.21192)]
62. Goffinet S, Schmitz D, Zavalysyn I, Legay A, Riviere E. Controlling security rules using natural dialogue: an application to smart home care. In: *Adjunct Proceedings of the 2021 ACM International Joint Conference on Pervasive and Ubiquitous Computing and Proceedings of the 2021 ACM International Symposium on Wearable Computers*. 2021 Presented at: UbiComp '21; September 21-26, 2021; Virtual p. 194-197. [doi: [10.1145/3460418.3479331](https://doi.org/10.1145/3460418.3479331)]
63. Alrumayh AS, Lehman SM, Tan CC. Context aware access control for home voice assistant in multi-occupant homes. *Pervasive Mob Comput* 2020 Sep;67:101196. [doi: [10.1016/j.pmcj.2020.101196](https://doi.org/10.1016/j.pmcj.2020.101196)]

Abbreviations

PSW: personal support worker

TAM: technology acceptance model

VA: voice assistant

Edited by A Kushniruk; submitted 03.03.22; peer-reviewed by Y Kumah-Crystal, D Chrimes; comments to author 01.05.22; revised version received 11.05.22; accepted 31.05.22; published 30.06.22.

Please cite as:

Tennant R, Allana S, Mercer K, Burns CM

Caregiver Expectations of Interfacing With Voice Assistants to Support Complex Home Care: Mixed Methods Study

JMIR Hum Factors 2022;9(2):e37688

URL: <https://humanfactors.jmir.org/2022/2/e37688>

doi: [10.2196/37688](https://doi.org/10.2196/37688)

PMID: [35771594](https://pubmed.ncbi.nlm.nih.gov/35771594/)

©Ryan Tennant, Sana Allana, Kate Mercer, Catherine M Burns. Originally published in *JMIR Human Factors* (<https://humanfactors.jmir.org>), 30.06.2022. This is an open-access article distributed under the terms of the Creative Commons Attribution License (<https://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in *JMIR Human Factors*, is properly cited. The complete bibliographic information, a link to the original publication on <https://humanfactors.jmir.org>, as well as this copyright and license information must be included.

Original Paper

Interdisciplinary Collaborations in Digital Health Research: Mixed Methods Case Study

Grit Krause-Jüttler¹, Dr; Jürgen Weitz^{1,2}, Prof Dr; Ulrich Bork¹, PD Dr

¹Department of Visceral, Thoracic and Vascular Surgery, Faculty of Medicine, Technische Universität Dresden, University Hospital Carl Gustav Carus, Dresden, Germany

²National Center for Tumor Diseases Dresden (NCT/UCC), German Cancer Research Center (DKFZ), Heidelberg, Helmholtz-Zentrum Dresden - Rossendorf (HZDR), Dresden, Germany

Corresponding Author:

Ulrich Bork, PD Dr

Department of Visceral, Thoracic and Vascular Surgery

Faculty of Medicine, Technische Universität Dresden

University Hospital Carl Gustav Carus

Fetscherstrasse 74

Dresden, 01307

Germany

Phone: 49 351458 ext 2742

Email: Ulrich.Bork@uniklinikum-dresden.de

Abstract

Background: Digital innovations in medicine are disruptive technologies that can change the way diagnostic procedures and treatments are delivered. Such innovations are typically designed in teams with different disciplinary backgrounds. This paper concentrates on 2 interdisciplinary research teams with 20 members from the medicine and engineering sciences working jointly on digital health solutions.

Objective: The aim of this paper was to identify factors on the individual, team, and organizational levels that influence the implementation of interdisciplinary research projects elaborating on digital applications for medicine and, based on the results, to draw conclusions for the proactive design of the interdisciplinary research process to make these projects successful.

Methods: To achieve this aim, 2 interdisciplinary research teams were observed, and a small case study (response rate: 15/20, 75%) was conducted using a web-based questionnaire containing both closed and open self-report questions. The Spearman rank correlation coefficient was calculated to analyze the quantitative data. The answers to the open-ended questions were subjected to qualitative content analysis.

Results: With regard to the interdisciplinary research projects investigated, the influencing factors of the three levels presented (individual, team, and organization) have proven to be relevant for interdisciplinary research cooperation.

Conclusions: With regard to recommendations for the future design of interdisciplinary cooperation, management aspects are addressed, that is, the installation of a coordinator, systematic definition of goals, required resources, and necessary efforts on the part of the involved interdisciplinary research partners. As only small groups were investigated, further research in this field is necessary to derive more general recommendations for interdisciplinary research teams.

Trial Registration: German Clinical Trials Register, DRKS00023909, https://www.drks.de/drks_web/navigate.do?navigationId=trial.HTML&TRIAL_ID=DRKS00023909 ; German Clinical Trials Register, DRKS00025077, https://www.drks.de/drks_web/navigate.do?navigationId=trial.HTML&TRIAL_ID=DRKS00025077

(*JMIR Hum Factors* 2022;9(2):e36579) doi:[10.2196/36579](https://doi.org/10.2196/36579)

KEYWORDS

team science; interdisciplinary; research collaboration; digital health; team processes

Introduction

Background

Digital innovations in medicine are disruptive technologies that can change the way diagnostic procedures and treatments are delivered. Innovation is usually designed in teams with different disciplinary backgrounds. Collaboration between professionals and experts from different educational backgrounds can release creative energies [1]. However, for successful interdisciplinary cooperation, certain basic principles must be observed. At present, projects require the cooperation of people with different knowledge and skills, who together consider complex problems from interdisciplinary perspectives and pursue new paths. Interdisciplinary cooperation is particularly necessary for the development of new technical achievements such as health care technologies. However, a lack of knowledge about the skills of people from other disciplines and different *languages* and *cultures* leads to problems in the process of interdisciplinary research collaboration.

In contrast to projects where people from different disciplines work in different areas, that is, have different individual goals, people in interdisciplinary teams have goals that they can accomplish only if they work together. Therefore, interdisciplinarity refers to the mutual dependence of disciplines. To achieve common goals, procedures or methods must be *negotiated* between disciplines. If this can be achieved without conflict, existing methods can be improved or new activities can be created. It is important that team members are aware that goals are to be achieved together, which the individual disciplines cannot achieve. People from other disciplines must be granted skills, and *other opinions* must be taken seriously. All participants must be aware of their role in the team and organization, have respect for other disciplines, and see the common goal as their distinct goal.

However, this effort is worthwhile when innovative solutions for complex challenges arise by combining the strengths of all participants.

This paper concentrates on 2 interdisciplinary research projects involving 20 researchers working jointly at the interface of medicine and engineering sciences. Using sociological methods of qualitative and quantitative surveys, this study examined which factors influence the implementation of interdisciplinary research between medicine and engineering sciences and which approaches exist to successfully shape this form of collaborative research in the future.

The first interdisciplinary project considered is the ARAILIS (Augmented Reality and Artificial Intelligence Supported Laparoscopic Imagery in Surgery) project, which aims to develop a prototype for innovative computer-assisted surgery using augmented reality and artificial intelligence. It is designed to support surgeons in making decisions that increase accuracy and therefore reduce the likelihood of complications during liver surgery.

The second project is the interdisciplinary PROSPER (Platform for Operation Scheduling and Prediction Using Machine Learning) project, which aims to develop a platform that enables

efficient and data-based decision-making for operating room (OR) planning processes through machine learning and the use of artificial intelligence. Using retrospective OR data and expert knowledge modeling, an automated solution is created that precisely predicts surgery duration, guarantees continuous planning adaptation, and enables day-based, flexible planning of all surgeries in multiple ORs for optimal resource use and OR efficiency.

In both projects, experts from medicine, computer, and further engineering sciences are researching interdisciplinarily.

The reasons for involving different scientific disciplines in solving medical research questions are multifaceted. The advantage of interdisciplinary collaboration is seen above all in the fact that a multidisciplinary approach to so-called *real-world problems* delivers more reliable results that are closer to application. In addition, especially when dealing with complex problems, such as digitalization in the health care system, additional expertise that is not available per se in the medical field is required [2,3].

In this study, both projects were jointly investigated, because they have comparable characteristics. In both projects, an interdisciplinary collaboration among engineers, computer scientists, and surgeons takes place, and they are working on a similar topic dealing with artificial intelligence to support the decision-making of surgeons. Therefore, the study team decided to investigate them together to obtain a larger data basis for describing and analyzing collaboration processes between these different disciplines. According to the previous project descriptions, differences can only be found in the concrete results that the projects are focusing on (surgery planning platform, image-based assistance system), but the way of collaboration; the different disciplines involved; and their way of cooperation are comparable in both considered projects.

Brief Overview About the State of Research

Over the past 30 years, researchers have extensively dealt with the issues of teamwork and the cross-disciplinary composition of research teams. In this context, a definitional issue must first be addressed.

As Aboelela et al [4] pointed out on the basis of a literature review, previous research has found various forms of cross-disciplinary research collaboration. Their review explains that the forms described in the literature can be defined along a continuum in terms of the “quality of the actual integration of different disciplines,” the “degree of cooperation (interaction of the researchers involved, communication and exchange of information),” and regards the “outcome of the collaboration”; that is, a concretely achieved solution [4]. On the basis of their research results, the authors distinguish the concepts of “multidisciplinarity,” “interdisciplinarity,” and “transdisciplinarity” with respect to the characteristics “participants/discipline,” “problem definition,” “research style,” and “presentation of findings” [4].

In accordance with this review, for the purpose of this paper, interdisciplinary research is understood as follows:

any study or group of studies undertaken by scholars from two or more distinct scientific disciplines. The research is based upon a conceptual model that links or integrates theoretical frameworks from those disciplines, uses study design and methodology that is not limited to any one field, and requires the use of perspectives and skills of the involved disciplines through-out multiple phases of the research process [4]

So, in contrast to *multidisciplinarity*, it is not a matter of additive cooperation in which disciplines work on partial aspects and develop their own solutions in parallel, but rather of a mutual expansion and integration of methods and solution approaches and thus a mutual compensation of existing gaps in the respective discipline with regard to the problem to be solved [5,6].

Previous research has dealt with the topic of research work in cross-disciplinary teams under the umbrella term *team science*. Klein [7] provides a brief overview of its various strands and distinguishes the following three main research clusters: *international network of interdisciplinary research* [8], the *transdisciplinary team science* (TTS) [9-11], and the *transdisciplinary trans-sector, problem-oriented research with stakeholders in society* localized in Europe [12]. Although the work of the first and third research clusters does not focus on any research area per se, TTS focuses on the field of interdisciplinary medical research and its cooperation with other, nonmedical scientific disciplines with the aim of answering complex questions, such as the management of cancer or the digitalization of health care. Therefore, the (TTS) available research results proved to be particularly relevant to the results presented in this paper. For completeness, reference should be made to the field of *Interprofessional Health Practice and Education* [13], which focuses on interprofessional cooperation in medical care.

Generally, both TTS and Interprofessional Health Practice and Education deal with the questions of how social factors influence interdisciplinary collaboration and how collaboration must be organized to work successfully in interdisciplinary teams [11] in the aforementioned sense [4]. Compared with disciplinary research, interdisciplinary collaboration poses some challenges. To make individuals from different disciplinary backgrounds and with different organizational affiliations collaborate successfully, an increased effort for communication and a high investment of time are required, especially in the early stages of collaboration. These investments are necessary to develop a common understanding of the research question, to make the different objectives of the participants involved in the research project transparent, and to establish an understanding of the respective conditions in the participating organizations [14]. Therefore, interdisciplinarity was not a success. Numerous influencing factors at various levels promote the success of interdisciplinary research projects. Interdisciplinarity can only cause a real benefit if these influencing factors are known and considered when coordinating cooperation [15-17].

Publications available to date have identified the relevant factors that influence the success of interdisciplinary cooperation in

the *individuals* involved, in the interaction within the *team*, and in the conditions for interdisciplinary research within the involved *organizations* [11,13,18].

However, successful interdisciplinarity has often been mentioned, but how can it be defined and measured? Tigges et al [14] provide an initial overview. There are 2 forms of measuring the success of interdisciplinary research work: first, the quantitative counting of results, such as publications or the amount of acquired external funds. Second, the most common type is the use of preformulated items for self-report to determine the individual perception of the involved researchers regarding the results of interdisciplinary research. Various instruments for the subjective assessment of the research process and the quality of interdisciplinary collaboration already exist, but have not yet been standardized. Such instruments can be found in various publications [18-24].

Research Questions

The introduced ARAILIS and PROSPER projects can only achieve the planned research and development objectives if they implement successful cooperation between representatives from the disciplines of medicine, computer, and engineering sciences. On the basis of the briefly outlined state of research regarding interdisciplinary research cooperation in the medical context, this paper focuses on the *first question*, which relation exists between the individual attitudes of the researchers involved, their perceptions of the cooperative research process at the team level, organizational framework conditions, and functioning interdisciplinary research cooperation.

As a *second question*, this paper focuses on recommendations for shaping interdisciplinary research collaborations between medicine and engineering sciences. The goal is to draw conclusions from the results for the proactive design of the interdisciplinary research process and thus ensure the achievement of projects' defined technical objectives with special attention to the maybe specific situation of interdisciplinary research in the medical context.

Methods

Overview

To answer these research questions, we conducted a small case study to investigate the introduced interdisciplinary projects. The study took place through a web-based questionnaire using the Lime Survey tool containing both closed and open questions for self-reporting. This study was conducted according to the process described in [Figure 1](#).

For quantitative questions, we mainly used existing measurement methods [18-24]. The operationalization for measuring the different variables displayed in [Figure 2](#), including their respective literature sources, is presented in [Multimedia Appendix 1](#) [20,23]. With the use of open questions in the web-based questionnaire, the respondents were asked to formulate with their own words their current impressions about the implementation and design of interdisciplinary research work in the observed projects.

Figure 1. Research process of this study.

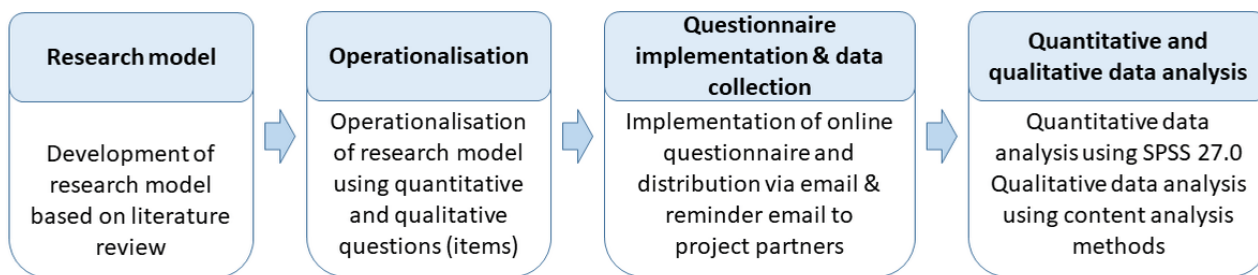
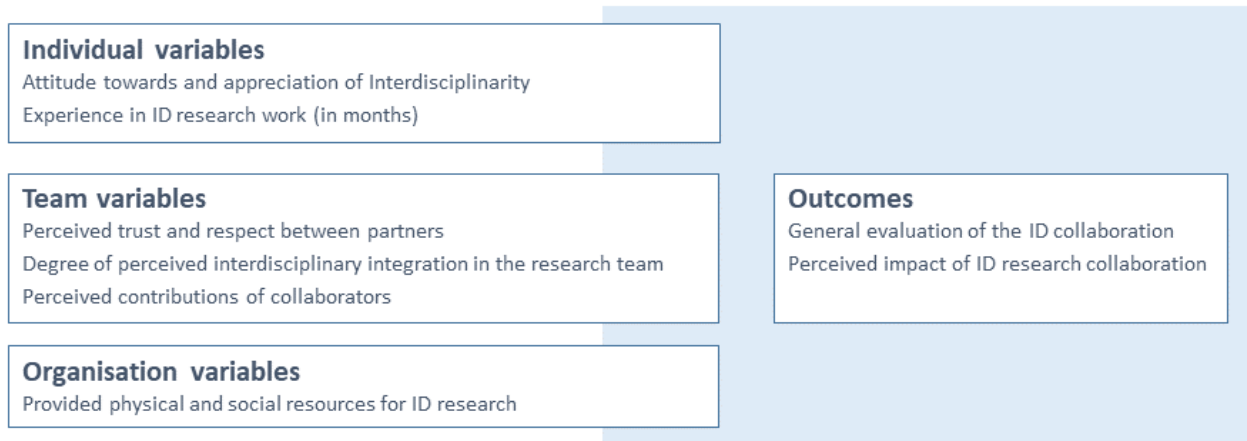


Figure 2. Research model for investigating variables that may be associated with a functioning interdisciplinary (ID) research collaboration.



Ethics Approval

No ethics committee approval was obtained for this anonymous survey trial in accordance with our internal review board guidelines (IRB00001473).

Recruitment

In these 2 interdisciplinary projects, 20 researchers from different disciplines were involved. The participants were requested to complete the questionnaire. The survey was conducted in November and December 2020, when both projects were in the first year of their planned 3-year duration, meaning that some of the teams had not been working together for a long period. This investigation was planned as the first measurement point of a longitudinal study with at least one additional measurement point to observe the development of social interaction processes of project members over time in relation to the final project results. The links providing access to the web-based questionnaire were sent via email to define the final date. Two weeks after the first email invitation, all potential survey participants received a reminder. Finally, 15 (in some cases 16) responded to the survey. Studies based on data collected in web-based surveys from individuals normally reach

response rates of approximately 53% (SD 20.4%) [25]. Therefore, in this study, it is assumed that a response rate of 75% allows reliable conclusions to be drawn for the project teams under investigation.

For quantitative data analysis, the Spearman rank correlation coefficient (SRCC) was calculated using SPSS (version 27.0; IBM Corporation). The SRCC is a nonparametric measure that can be applied to very small sample sizes [26]. The correlation coefficient was calculated at the item level. The results are presented in Tables 1-3, which show selected descriptive measures (mean and SD) as well as the calculation results for significant correlations of items that were used to operationalize the individual, team, and organizational variables and the defined outcomes listed in Figure 2. Owing to the rather small sample size, only very close relationships between the variables examined proved to be significant [26]. All items were rated by 15 of the 20 project members contacted. Items assessed by 16 of the 20 are marked in Tables 1-3.

The answers to the open questions were subjected to qualitative content analysis based on the approach provided by Mayring [27] to abstract and summarize their essential content.

Table 1. Mean, SD, Spearman rank correlation coefficient (SRCC; 1-sided hypothesis test), and *P* values for individual variables and outcomes.

	Value, mean (SD)	Productivity of collaboration meetings	Overall productivity of collaboration	Integration of results succeeds well	Develop- ment of com- mon lan- guage suc- ceeds well	Develop- ment of a common theoretical basis suc- ceeds well	Our project team is suc- cessful	My sub- project is success- ful	In general, collaboration has im- proved my research pro- ductivity
Outcomes									
Productivity of collaboration meetings (n=16)	3.13 (0.72)	— ^a	—	—	—	—	—	—	—
SRCC									
<i>P</i> value									
Overall productivity of collaboration (n=16)	3 (81) (0.66)		—	—	—	—	—	—	—
SRCC		0.30							
<i>P</i> value		.13							
In our project team									
Integration of results succeeds well	3.67 (0.82)			—	—	—	—	—	—
SRCC		-0.01	0.47						
<i>P</i> value		.48	.04						
Development of common language succeeds well	3.73 (0.59)				—	—	—	—	—
SRCC		0.28	0.48	0.67					
<i>P</i> value		.16	.03	.003					
Development of a common theoretical basis succeeds well	3.60 (0.74)					—	—	—	—
SRCC		-0.06	0.16	0.51	0.38				
<i>P</i> value		.41	.28	.03	.08				
Measured by results so far									
Our project team is successful	3.67 (0.82)						—	—	—
SRCC		0.05	0.61	0.66	0.39	0.30			
<i>P</i> value		.43	.008	.003	.07	.14			
My subproject is successful	3.73 (0.80)							—	—
SRCC		0.06	0.48	0.60	0.29	0.20	0.93		
<i>P</i> value		.41	.04	.01	.15	.23	<.001		
In general, collaboration has improved my research productivity	4.07 (0.70)								—
SRCC		-0.06	-0.02	0.43	0.38	0.74	0.20	0.20	
<i>P</i> value		.42	.47	.05	.08	.001	.24	.24	
Individual variables									

	Value, mean (SD)	Productivity of collaboration meetings	Overall productivity of collaboration	Integration of results succeeds well	Development of common language succeeds well	Development of a common theoretical basis succeeds well	Our project team is successful	My sub-project is successful	In general, collaboration has improved my research productivity
I am optimistic that ID^b research among project collaborators will lead to valuable scientific outcomes that would not have occurred without that kind of collaboration^c (n=16)	4.63 (0.50)								
SRCC		-0.17	0.35	0.20	0.07	0.26	0.51	0.40	-0.12
<i>P</i> value		.26	.09	.23	.40	.17	.03	.07	.33
Participating in an ID team improves the results that are developed^c	4.20 (0.86)								
SRCC		-0.16	0.23	0.37	0.06	-0.00	0.32	0.44	-0.13
<i>P</i> value		.28	.21	.09	.42	.50	.13	.05	.33

^aNot applicable.

^bID: interdisciplinary.

^cWhat do you think about the interdisciplinary research process in the project from your individual point of view? Please rate your views using gradations 1 "strongly disagree," 2 "somewhat disagree," 3 "not sure," 4 "somewhat agree," and 5 "strongly agree."

Table 2. Mean, SD, Spearman rank correlation coefficient (SRCC; 1-sided hypothesis test), and *P* values for team variables and outcomes.

	Value, mean (SD)	Productivity of collaboration meetings	Overall Productivity of collaboration	Integration of results succeeds well	Develop- ment of com- mon lan- guage suc- ceeds well	Develop- ment of a common theoretical basis suc- ceeds well	Our project team is suc- cessful	My sub- project is success- ful	In general, collaboration has im- proved my research pro- ductivity
Outcomes									
Productivity of collaboration meetings (n=16)	3.13 (0.72)	— ^a	—	—	—	—	—	—	—
SRCC									
<i>P</i> value									
Overall Productivity of collaboration (n=16)	3.81 (0.66)		—	—	—	—	—	—	—
SRCC		0.30							
<i>P</i> value		.13							
In our project team									
Integration of results succeeds well	3.67 (0.82)			—	—	—	—	—	—
SRCC		-0.01	0.47						
<i>P</i> value		.48	.04						
Development of common language succeeds well	3.73 (0.59)				—	—	—	—	—
SRCC		0.28	0.48	0.67					
<i>P</i> value		.16	.03	.003					
Development of a common theoretical basis succeeds well	3.60 (0.74)					—	—	—	—
SRCC		-0.06	0.16	0.51	0.38				
<i>P</i> value		.41	.28	.03	.08				
Measured by results so far									
Our project team is successful	3.67 (0.82)						—	—	—
SRCC		0.05	0.61	0.66	0.39	0.30			
<i>P</i> value		.43	.008	.003	.07	.14			
My subproject is successful	3.73 (0.80)							—	—
SRCC		0.06	0.48	0.60	0.29	0.20	0.93		
<i>P</i> value		.41	.04	.01	.15	.23	<.001		
In general, collaboration has improved my research productivity	4.07 (0.70)								—
SRCC		-0.06	-0.02	0.43	0.38	0.74	0.20	0.20	
<i>P</i> value		.42	.42	.05	.08	.001	.24	.24	
Team variables									

	Value, mean (SD)	Productivity of collaboration meetings	Overall Productivity of collaboration	Integration of results succeeds well	Development of common language succeeds well	Development of a common theoretical basis succeeds well	Our project team is successful	My sub-project is successful	In general, collaboration has improved my research productivity
Acceptance of new ideas^b (n=16)	4.19 (0.91)								
SRCC		0.50	0.43	0.28	0.39	0.57	0.49	0.38	0.21
<i>P</i> value		.02	.048	.15	.07	.01	.03	.08	.23
Communication among collaborators^b (n=16)	3.88 (0.62)								
SRCC		0.55	0.42	0.24	0.15	0.11	0.41	0.27	0.00
<i>P</i> value		.01	.06	.20	.30	.35	.06	.16	.50
Resolution of conflicts among collaborators^b (n=16)	3.94 (0.57)								
SRCC		0.22	0.18	-0.01	0.32	0.44	0.25	0.10	0.17
<i>P</i> value		.21	.25	.48	.13	.05	.19	.36	.27
Ability to accommodate different working styles of collaborators^b (n=16)	4.13 (0.81)								
SRCC		0.52	0.53	0.30	0.52	0.20	0.52	0.37	0.05
<i>P</i> value		.02	.02	.14	.02	.24	.02	.09	.43
Integration of research methods from different fields^b (n=16)	3.69 (0.70)								
SRCC		-0.02	0.38	0.67	0.23	0.40	0.67	0.51	0.19
<i>P</i> value		.47	.08	.003	.21	.07	.003	.03	.25
Integration of theories and models from different fields^b(n=16)	3.63 (0.72)								
SRCC		-0.01	0.34	0.76	0.37	0.50	0.59	0.43	0.19
<i>P</i> value		.48	.10	.001	.08	.03	.01	.05	.25
Involvement of collaborators from diverse disciplines^b (n=16)	4.19 (0.66)								
SRCC		-0.23	0.47	0.55	0.48	0.51	0.62	0.46	0.12
<i>P</i> value		.20	.03	.02	.04	.03	.007	.04	.34
High motivation for collaboration^c	4.40 (0.51)								
SRCC		0.11	0.49	0.31	0.36	0.45	0.31	0.24	0.52
<i>P</i> value		.34	.03	.13	.09	.045	.13	.19	.02

	Value, mean (SD)	Productivity of collaboration meetings	Overall Productivity of collaboration	Integration of results succeeds well	Development of common language succeeds well	Development of a common theoretical basis succeeds well	Our project team is successful	My sub-project is successful	In general, collaboration has improved my research productivity
Reliable fulfillment of tasks taken over within the project team^c	3.93 (0.60)								
SRCC		0.40	0.71	0.51	0.38	0.11	0.43	0.29	0.17
<i>P</i> value		.0	.002	.03	.08	.35	.06	.15	.27
Willingness to coordinate one's own research work with the others in the project team and to work intensively with the other project members^c	4.27 (0.80)								
SRCC		-0.04	0.26	0.00	-0.04	0.49	0.27	0.13	0.58
<i>P</i> value		.44	.17	.50	.45	.03	.16	.32	.01
Interest in other disciplines involved and willingness to recognize other disciplines as equivalent^c	4.33 (0.62)								
SRCC		-0.09	0.27	0.53	0.66	0.49	0.35	0.19	0.66
<i>P</i> value		.38	.16	.02	.004	.03	.10	.25	.004

^aNot applicable.

^bWhen thinking about the researchers collaborating on the project, how do you evaluate the following aspects? Please use the gradation 1 "inadequate," 2 "poor," 3 "satisfactory," 4 "good," or 5 "excellent"!

^cWith regard to your experiences in the project so far, what impressions do you have regarding the research contributions of your collaborators? How do the following statements apply: Please evaluate the mentioned issues using gradation 1 "does not apply at all," 2 "does more likely apply," 3 "does partly apply," 4 "does more likely apply," or 5 "does strongly apply."

Table 3. Mean, SD, Spearman rank correlation coefficient (SRCC; 1-sided hypothesis test), and *P* values for organization variables and outcomes.

	Value, mean (SD)	Productivity of collaboration meetings	Overall productivity of collaboration	Integration of results succeeds well	Development of common language succeeds well	Development of a common theoretical basis succeeds well	Our project team is successful	My subproject is successful	In general, collaboration has improved my research productivity
Outcomes									
Productivity of collaboration meetings (n=16)	3.13 (0.72)	— ^a	—	—	—	—	—	—	—
SRCC									
<i>P</i> value									
Overall productivity of collaboration (n=16)	3.81 (0.66)		—	—	—	—	—	—	—
SRCC		0.30							
<i>P</i> value		.13							
In our project team									
Integration of results succeeds well	3.67 (0.82)			—	—	—	—	—	—
SRCC		−0.01	0.47						
<i>P</i> value		.48	.04						
Development of common language succeeds well	3.73 (0.59)				—	—	—	—	—
SRCC		0.28	0.48	0.67					
<i>P</i> value		.16	.03	.003					
Development of a common theoretical basis succeeds well	3.60 (0.74)					—	—	—	—
SRCC		−0.06	0.16	0.51	0.38				
<i>P</i> value		.41	.28	.03	.08				
Measured by results so far									
Our project team is successful	3.67 (0.82)						—	—	—
SRCC		0.05	0.61	0.66	0.39	0.30			
<i>P</i> value		.43	.008	.003	.07	.14			
My subproject is successful	3.73 (0.80)							—	—
SRCC		0.06	0.48	0.60	0.29	0.20	0.93		
<i>P</i> value		.41	.04	.01	.15	.23	<.001		
In general, collaboration has improved my research productivity	4.07 (0.70)								—
SRCC		−0.06	−0.02	0.43	0.38	0.74	0.20	0.20	
<i>P</i> value		.42	.47	.05	.08	.001	.24	.24	
Organization variables									

	Value, mean (SD)	Productivity of collaboration meetings	Overall productivity of collaboration	Integration of results succeeds well	Development of common language succeeds well	Development of a common theoretical basis succeeds well	Our project team is successful	My sub-project is successful	In general, collaboration has improved my research productivity
Physical resources for ID^b research: availability of physical space (eg, office, laboratory etc)^c	3.80 (0.94)								
SRCC		0.23	0.45	0.37	0.13	0.46	0.46	0.53	0.13
<i>P</i> value		.20	.045	.09	.32	.04	.04	.02	.32
Physical resources for ID research: availability of electronic or other resources for collaboration between remote research sites (knowledge management systems, online platforms and cloud services, etc)^c	3.93 (1.22)								
SRCC		0.61	0.10	-0.06	-0.21	-0.21	-0.11	0.01	-0.29
<i>P</i> value		.008	.37	.42	.23	.23	.35	.49	.14
Social resources for ID research: my involvement in an ID research project is highly appreciated by my supervisors^c	4.07 (0.59)								
SRCC		0.55	0.31	0.02	0.28	0.07	0.37	0.38	0.15
<i>P</i> value		.02	.13	.47	.16	.40	.09	.08	.29
Social resources for ID research: my involvement in an ID research project is highly appreciated by my colleagues^c	3.73 (0.70)								
SRCC		0.33	0.54	0.57	0.55	0.31	0.74	0.64	0.32
<i>P</i> value		.11	.02	.01	.02	.13	.001	.005	.12

^aNot applicable.

^bID: interdisciplinary.

^cConsidering the provided institutional or social resources for conducting the interdisciplinary research work in the project, how do you evaluate the availability of the following issues? For your evaluation, please use the gradation 1 "inadequate," 2 "poor," 3 "satisfactory," 4 "good," or 5 "excellent"!

Collection of Quantitative Data

In accordance with the state of research briefly presented earlier, the closed survey questions focused on the subjective assessment of the individual, team, and organization variables that might be associated with a functioning interdisciplinary research collaboration.

Figure 2 lists the variables collected. Multimedia Appendix 1 contains their concrete operationalization in the survey questionnaire based on existing instruments [19-22] and the development of their own measurements.

Individual variables measure the respondent's personal attitude toward and evaluation of interdisciplinary research collaboration as well as the respondent's own experience with it. It is assumed that a positive attitude and long-term interdisciplinary experience

are associated with positive perceptions of interdisciplinary outcomes.

Team variables cover the aspects of perceived trust between partners, the self-assessed degree of interdisciplinary integration in the research team, and the extent to which partners contribute to team outcomes. It is assumed that the positive ratings of these variables are associated with the positive ratings of interdisciplinary outcomes.

Organization variables refer to the support the respondents perceive for interdisciplinary research within their home organizations; for example, whether the necessary physical resources and social support from supervisors and colleagues are available for this.

Outcomes were measured by means of a general individual assessment of the results achieved so far and the productivity of the interdisciplinary research team. The collection of objective key figures, such as the number of publications, did not yield any results at this early point in the projects and will be collected again at later measurement points.

Collection of Qualitative Data

Furthermore, the questionnaire included open-ended questions. Respondents answered the following self-developed questions (the sample size is provided in brackets):

1. What are the advantages and the disadvantages of interdisciplinary research work? (n=14)
2. Which facilitators and barriers for interdisciplinary research work do you personally perceive? Please think about aspects on individual, team, and organizational levels (eg, professorship or institute). (n=11)
3. Based on your experiences in interdisciplinary research so far, do you perceive differences between the collaboration with physicians and the collaboration with other disciplines? If yes, which differences do you see? (n=14)
4. When you recap your experiences in interdisciplinary research so far, which recommendations would you give to the project team and beyond to make interdisciplinary research collaborations successful? (n=13)

Results

A demographic description of the surveyed project member sample is not provided because of the small group investigated and the possibility of an individual identification of responders; no demographic data of the project members were collected.

Quantitative Results of the Correlation Analyses

Table 1 shows few significant results for the individual variables. Respondents who are very strongly optimistic that interdisciplinary will lead to valuable research results that would not have been produced otherwise also strongly agree that the project is successful (SRCC=0.51; $P=.03$). Respondents who strongly believe that interdisciplinary will improve the research results produced also strongly believe that their own subproject is successful (SRCC=0.44; $P=.05$). A significant relationship between the length of personal experience with interdisciplinary research and outcomes is not found for the studied sample and, therefore, not mentioned in the correlation table.

Numerous significant relationships emerged with the defined outcomes for the *team variables* listed in Table 2. Strong correlations (SRCC \geq 0.50) were shown for respondents who rated the acceptance of new ideas in the team as high. These respondents simultaneously perceived the high productivity of project meetings (SRCC=0.50; $P=.02$) and successful development of a common theoretical base (SRCC=0.57; $P=.01$). The perception of excellent communication in the interdisciplinary team was strongly associated with the productivity of project meetings rated high (SRCC=0.55; $P=.01$).

Individuals who rated the interdisciplinary team as very good at reconciling the different work styles of the collaboration partners, rated the productivity of work meetings (SRCC=0.52; $P=.02$), the overall productivity of the collaboration (SRCC=0.53; $P=.02$), the development of a common language (SRCC=0.52; $P=.02$), and, as measured by previous results, the success of the project team (SRCC=0.52; $P=.02$) as very high. The finding of a very strong integration of research methods from different fields is very closely related to the perception of a very high integration of results (SRCC=0.67; $P=.003$), successful work on the level of the overall project (SRCC=0.67; $P=.003$), and the respective subproject (SRCC=0.51; $P=.03$). The perception of a very successful integration of theories and models from different research fields is again strongly correlated with a high assessment of a successful integration of results (SRCC=0.76; $P=.001$), the perception of successful development of a collaborative theoretical basis (SRCC=0.50; $P=.03$), and successful project implementation (SRCC=0.59; $P=.001$). Respondents who assessed the involvement of collaborative partners from different disciplines as very good, rated the integration of results (SRCC=0.55; $P=.02$), the development of a collaborative theoretical base (SRCC=0.51; $P=.03$) as very high, and the project team as very successful (SRCC=0.62; $P=.007$). The perception of high motivation for collaboration among team members was significantly positively correlated with the perception that collaboration increased one's research productivity (SRCC=0.52; $P=.02$). The perception that tasks taken on in the interdisciplinary team are reliably completed by team colleagues is closely related to the positive assessment of collaboration productivity (SRCC=0.71; $P=.002$) and to a perceived very successful integration of results in the team (SRCC=0.51; $P=.03$). An existing willingness to coordinate one's own research with team members and to collaborate intensively with members is strongly associated with the assessment that interdisciplinary collaboration also greatly improves one's research productivity (SRCC=0.58; $P=.01$). The impression that team members show interest in the disciplines involved and are willing to perceive them as equals is strongly associated with a perceived successful integration of results (SRCC=0.53; $P=.02$) and is strongly associated with the development of a common language (SRCC=0.66; $P=.004$) and the assessment that collaboration has improved one's own research productivity (SRCC=0.66; $P=.004$).

The results for the *organization variables* contained in Table 3 point to the high relevance of physical resources (eg, offices and laboratories) for interdisciplinary research. In this case, significant correlations were found between the positive evaluation of their availability and collaboration productivity

(SRCC=0.45; $P=.045$), the development of a common theoretical basis (SRCC=0.46; $P=.04$), a project perceived as successful (SRCC=0.46; $P=.04$), and a subproject perceived as successful (SRCC=0.53; $P=.02$) in the context of interdisciplinary collaboration. A positive evaluation of the availability of electronic resources for location-independent collaboration (knowledge management systems, web-based platforms, etc) shows a strong correlation with the assessment of a high productivity of meetings (SRCC=0.61; $P=.008$). As a side note, it should be added that perceived low availability of these electronic resources for collaboration shows a negative relationship with several aspects of outcome integration and success evaluation, although these are not significant. The availability of social resources is also relevant for outcome evaluation. High perceived supervisor support for interdisciplinary collaboration was significantly related to the highly rated productivity of project meetings (SRCC=0.55; $P=.02$). However, social support from colleagues in interdisciplinary research collaboration appears to be of greater importance. Respondents who perceived high appreciation by colleagues for their own involvement in interdisciplinary research simultaneously rated the productivity of collaboration highly (SRCC=0.54; $P=.02$), perceived strong integration of results in the interdisciplinary team (SRCC=0.57; $P=.01$), were more likely to rate the development of a common language as successful (SRCC=0.55; $P=.02$), and perceived both the overall project (SRCC=0.74; $P=.001$) and their own subproject (SRCC=0.64; $P=.005$) as more successful.

Qualitative Results of the Open Questions

Textbox 1 summarizes the answers to question 1 regarding the advantages and disadvantages of interdisciplinary research.

These advantages are mainly seen in gaining new knowledge and methods from other disciplines. The respondents saw disadvantages in terms of the necessity of providing more time.

The answers to *question 2* are summarized in **Textbox 2**. These are facilitating factors for interdisciplinary research work in the establishment of a functioning project management (project manager, coordination of tasks, creation of common goals) but also in individual factors, such as the willingness to engage with other disciplines and to accept them.

In response to *question 3* about differences in collaboration with physicians compared to other disciplines, 6 (43%) of the 14 respondents stated that they did not perceive any differences. The remaining respondents mainly reported that they perceived physicians to be heavily involved in clinical work and therefore had less time for joint research work. One of the interviewees described it in such a way that physicians are seen more as outside experts who are only contacted when necessary, whereas the remaining multidisciplinary research partners are perceived as equal team members. One respondent also mentioned the problem of high data protection requirements for the use of patient data and the associated hurdles as a difference in cooperation with physicians compared with other disciplines.

Finally, the respondents provided their *recommendations* for the successful design of interdisciplinary research collaborations, which are summarized in **Textbox 3**. They are predominantly emphasizing the need for systematic collaboration management.

Textbox 1. Summarized answers regarding advantages and disadvantages of interdisciplinary research work (open question 1).

Advantages of interdisciplinary research (in comparison to monodisciplinary research)

- Development of new ideas
- Knowledge transfer between different disciplines
- Leverage of different (technical) expertise and perspectives
- Broadening of mind and knowledge
- Experience of new methods
- Critical assessment of own methods, tools by other disciplines
- Solution of more complex problems by developing overall advanced systems

Disadvantages of interdisciplinary research (in comparison to monodisciplinary research)

- More time is needed for:
 - Providing and gaining information, explanation
 - Finding a common language
 - Building a mutual understanding

Textbox 2. Summarized answers regarding facilitators and barriers for interdisciplinary research work (open question 2).

Perceived facilitators for interdisciplinary research

- Supportive, reliable team members
- Independent project manager
- Good management or coordination by project leader
- Input from senior scientists
- Same or common goals
- Willingness to get involved in other scientific disciplines
- Acceptance that interdisciplinary needs more time

Perceived barriers for interdisciplinary research

- Different technical but also native languages (eg, different meanings of same terms)
- Dependency on work of others for own results
- Geographic distance between partners
- Focus on monodisciplinary research results
- Missing support for administrative issues
- Data protection problem regards medical data

Textbox 3. Summarized answers regarding recommendations for successful interdisciplinary research work (open question 4).

Recommendations for successful interdisciplinary research collaboration

- Concrete definition of project objectives, workload, and requirements
- Installation of a project coordinator
- Implementation of regular meetings with progress reports
- Flexibility for changes and adaptations in the workplan
- Clarification of roles and expected contributions for each single project member
- Establishment of a common language

Discussion

Principal Findings

With regard to the quantitative results for the interdisciplinary research projects discussed here, influencing factors of all 3 levels presented (individual, team, and organization) have proven to be relevant for a cooperation that is perceived as successful. Positive individual attitudes toward interdisciplinary research work are related to positive outcomes. Pre-existing experience with interdisciplinarity did not play a role in the sample investigated, although this has already been shown in other studies [28]. The results of the team variables indicate that well-functioning group processes, in the sense of mutual acceptance and real integration of theories, methods, and approaches, are reflected in the perceived results of interdisciplinary collaboration. Thus, the results of the analyzed organizational factors indicate the high relevance of the provision of physical and social resources for successful interdisciplinary research collaboration. The high relevance of electronic resources for remote collaboration may be due to COVID-19 pandemic-related social distancing measures.

The answers to the open, more qualitative-oriented questions describe the advantages of interdisciplinary research work in getting to know new ideas, methods, and knowledge and integrating them into their own work. They also identified an advantage in the fact that more complex problems are solved that require the inclusion of different professional perspectives. They observed disadvantages, particularly the fact that this form of cooperation is more time-consuming, as more explanations and the establishment of a common understanding are necessary. Regarding recommendations for the future design of interdisciplinary cooperation, management aspects are addressed here, ie, goals, required resources, and necessary efforts on the part of the involved interdisciplinary research partners should be clearly defined in advance. Among other things, establishing a project coordinator and holding regular meetings are recommended. Furthermore, social aspects, such as the definition of the roles of each individual participant in the entire team and the establishment of a common language; for example, clarified common terms, should also be considered.

Limitations

First, this study provides interesting insights into the projects introduced at the beginning of the article, but, at the same time,

these insights are mainly limited to both projects. It is possible to derive some general conclusions and recommendations that are also covered by the existing research literature and results [11,17], but they must be viewed with some caution.

A further limiting factor was the statistical calculations that were performed. For the correlation analyses, the SRCC was calculated, which is a nonparametric measure that can be applied to very small sample sizes. It must be considered that these correlations do not allow any statement regarding a causal effect but only give hints about which aspects could be associated.

Moreover, the results were based on the self-reports of the responding researchers. Although personal opinions and intentions are relevant for individual behavior, this limitation to only one data source prevents additional data validation (eg, comparison of self and external assessment). In this regard, aspects such as common method bias and common method variance are of interest because all variables are measured using the same instrument [29,30]. Another limitation is that the variables considered in the research model (Figure 2) were very selective. These were chosen based on literature research, but not all identified influencing factors that were part of the questionnaire study became part of the data analysis presented in this paper. This is because no significant relationship was confirmed. Furthermore, it must be assumed that there is content overlap between the different variables presented in the research model (Figure 2). Further development of these instruments is necessary.

A further fact for consideration is the limited generalizability due to the survey period, which was during the second lockdown in Germany, caused by measures implemented due to the COVID-19 pandemic. In this regard, the mode of cooperation was based on web-based tools (videoconferences and email), and personal encounters did not occur. This certainly influenced the response behavior.

Implications and Recommendations

Although the results are based on a small sample, it is possible to derive more general recommendations for the design and implementation of interdisciplinary research collaborations in digital health projects.

When providing recommendations for the design of successful interdisciplinary research collaboration, the levels of teams and organization can be considered because they are accessible to the direct influence of leadership and management. At the level of the concrete *team* that is collaborating, a project management regime should be implemented regarding the following aspects:

- Definition of a reliable and binding project plan including responsibilities, meetings, roles of all team members, timeline, and deadlines.
- Installation of a person who is and feels responsible for monitoring and complying with the plan (eg, project coordinator).
- Elaboration of a common understanding of the contents and objectives of the project.
- Establishment of team spirit and mutual trust as a precondition for openness and exchange of knowledge between research partners.

In addition, it must not be forgotten that the members of the interdisciplinary team are also members of the organizations where they are employed. At this *organizational level*, some aspects must be assured for successful collaboration to take place. As the results show, the following issues must be considered:

- Creation of an organizational atmosphere that demonstrates appreciation for interdisciplinary research work and is well-aware that it takes more time and effort in comparison to monodisciplinary research.
- Provision of social support (eg, recognition, affiliation, and instrumental assistance) by supervisors and colleagues for interdisciplinary research efforts.
- Deployment of an appropriate technical infrastructure that enables interdisciplinary collaborations even about spatial distance.

Regarding clinician scientists and their special role perceived as somehow external, organizational modes should be found to give them the opportunity for more integration in the whole interdisciplinary research team, which in turn will contribute to common understanding and trust and, therefore, positive results of the cooperative science process.

In this organizational context, *individual characteristics* could also be considered as far as personnel selection procedures are concerned. When a vacant position in an interdisciplinary research team has to be filled, potential recruits should be considered or selected that have already experience with interdisciplinary research or who at least seem to be open for that kind of cooperation and have the empathy to engage with other disciplines.

In the future, training courses that qualify team members and leaders for interdisciplinary research cooperation could be envisioned to enable them to act under the special conditions of interdisciplinary research. In this regard, more research is necessary [17,31].

Outlook

As mentioned at the beginning of this paper, interdisciplinary research, especially in the context of the digitalization of medicine and health care systems, has become increasingly important. To design and manage this kind of collaboration successfully, it is necessary to identify the *adjusting screws* at different levels in research organizations and beyond. Leaders, researchers, and students must be sensitized and trained for this type of cooperation. Constant research on these social respectively human factors influencing collaboration is essential, mainly regarding content-related aspects of training for interdisciplinary research [17,32]. However, from a methodological point of view, more sophisticated study designs for monitoring interdisciplinary research collaboration are necessary, especially regarding multivariate influences. In this regard, success indicators for interdisciplinary research should be extended beyond the dominance of bibliometrics [17]. In addition, for future research and bigger samples than that of this study, demographic data characterizing the actors involved should be collected to provide more information about the transferability of results to other fields of concern. Furthermore,

the concrete disciplinary composition of the respective project teams must be considered because the differences in working styles and professional cultures of single disciplines may also impact collaboration.

To some extent, the present case study shows that individual and team perceptions of success can also be used. Besides this,

already available and helpful results for managing interdisciplinary projects from social science disciplines should be integrated, reflecting topics such as *transformational leadership* and its impact on creativity [33] or team support roles [34].

Acknowledgments

The projects ARAILIS (Augmented Reality and Artificial Intelligence Supported Laparoscopic Imagery in Surgery) and PROSPER (Platform for Operation Scheduling and Prediction Using Machine Learning) are measures that are cofinanced with tax funds based on the budget passed by the Saxon State Parliament.

Conflicts of Interest

None declared.

Multimedia Appendix 1

Sources of used self-report items for collecting quantitative data.

[[DOCX File, 22 KB - humanfactors_v9i2e36579_app1.docx](#)]

References

1. Shin SJ, Zhou J. When is educational specialization heterogeneity related to creativity in research and development teams? Transformational leadership as a moderator. *J Appl Psychol* 2007 Nov;92(6):1709-1721. [doi: [10.1037/0021-9010.92.6.1709](https://doi.org/10.1037/0021-9010.92.6.1709)] [Medline: [18020807](https://pubmed.ncbi.nlm.nih.gov/18020807/)]
2. Choi BC, Pak AW. Multidisciplinarity, interdisciplinarity and transdisciplinarity in health research, services, education and policy: 1. Definitions, objectives, and evidence of effectiveness. *Clin Invest Med* 2006 Dec;29(6):351-364. [Medline: [17330451](https://pubmed.ncbi.nlm.nih.gov/17330451/)]
3. Shen L, Wang S, Dai W, Zhang Z. Detecting the interdisciplinary nature and topic hotspots of robotics in surgery: social network analysis and bibliometric study. *J Med Internet Res* 2019 Mar 26;21(3):e12625 [FREE Full text] [doi: [10.2196/12625](https://doi.org/10.2196/12625)] [Medline: [30912752](https://pubmed.ncbi.nlm.nih.gov/30912752/)]
4. Aboelela SW, Larson E, Bakken S, Carrasquillo O, Formicola A, Glied SA, et al. Defining interdisciplinary research: conclusions from a critical review of the literature. *Health Serv Res* 2007 Feb;42(1 Pt 1):329-346 [FREE Full text] [doi: [10.1111/j.1475-6773.2006.00621.x](https://doi.org/10.1111/j.1475-6773.2006.00621.x)] [Medline: [17355595](https://pubmed.ncbi.nlm.nih.gov/17355595/)]
5. Parthey H. Struktur wissenschaftlicher und gesellschaftlicher integrität von forschungssituationen. In: von Klaus Fischer H, Parthey H, editors. *Gesellschaftliche Integrität der Forschung: Wissenschaftsforschung Jahrbuch 2005*. Berlin, Germany: Gesellschaft für Wissenschaftsforschung; 2005:71-94.
6. Nicolescu B, Stavinschi M. *Transdisciplinarity in Science and Religion*. Bucharest, Romania: Curtea Veche Publishing House; 2007.
7. Klein JT. Evaluation of interdisciplinary and transdisciplinary research: a literature review. *Am J Prev Med* 2008 Aug;35(2 Suppl):S116-S123. [doi: [10.1016/j.amepre.2008.05.010](https://doi.org/10.1016/j.amepre.2008.05.010)] [Medline: [18619391](https://pubmed.ncbi.nlm.nih.gov/18619391/)]
8. Laudel G, Origi G. Introduction to a special issue on the assessment of interdisciplinary research. *Res Eval* 2006 Apr 01;15(1):2-4. [doi: [10.3152/147154406781776066](https://doi.org/10.3152/147154406781776066)]
9. Stokols D, Hall KL, Taylor BK, Moser RP. The science of team science: overview of the field and introduction to the supplement. *Am J Prev Med* 2008 Aug;35(2 Suppl):S77-S89. [doi: [10.1016/j.amepre.2008.05.002](https://doi.org/10.1016/j.amepre.2008.05.002)] [Medline: [18619407](https://pubmed.ncbi.nlm.nih.gov/18619407/)]
10. Lotrecchiano GR. A dynamical approach toward understanding mechanisms of team science: change, kinship, tension, and heritage in a transdisciplinary team. *Clin Transl Sci* 2013 Aug;6(4):267-278 [FREE Full text] [doi: [10.1111/cts.12054](https://doi.org/10.1111/cts.12054)] [Medline: [23919361](https://pubmed.ncbi.nlm.nih.gov/23919361/)]
11. O'Rourke M, Crowley S, Laursen B, Robinson B, Vasko SE. Disciplinary diversity in teams: integrative approaches from unidisciplinarity to transdisciplinarity. In: Hall KL, Vogel AL, Croyle RT, editors. *Strategies for Team Science Success: Handbook of Evidence-Based Principles for Cross-Disciplinary Science and Practical Lessons Learned from Health Researchers*. Cham, Switzerland: Springer; 2019:21-46.
12. Defila R, Di Gulio A. Managing consensus in inter- and transdisciplinary teams: tasks and expertise. In: Frodeman R, editor. *The Oxford Handbook of Interdisciplinarity*. 2nd ed. Oxford, UK: Oxford University Press; 2017:332-337.
13. McMurtry A, Kilgour K, Rohse S. Health research, practice, and education. In: Frodeman R, editor. *The Oxford Handbook of Interdisciplinarity*. 2nd ed. Oxford, UK: Oxford University Press; 2017.
14. Tigges BB, Miller D, Dudding KM, Balls-Berry JE, Borawski EA, Dave G, Measures of Collaboration Workgroup of the Collaboration and Engagement Domain Task Force, National Center for Advancing Translational Sciences, National

- Institutes of Health. Measuring quality and outcomes of research collaborations: an integrative review. *J Clin Transl Sci* 2019 Oct;3(5):261-289 [FREE Full text] [doi: [10.1017/cts.2019.402](https://doi.org/10.1017/cts.2019.402)] [Medline: [31660251](https://pubmed.ncbi.nlm.nih.gov/31660251/)]
15. Lynch J. It's not easy being interdisciplinary. *Int J Epidemiol* 2006 Oct;35(5):1119-1122. [doi: [10.1093/ije/dy1200](https://doi.org/10.1093/ije/dy1200)] [Medline: [16987842](https://pubmed.ncbi.nlm.nih.gov/16987842/)]
 16. Bozeman B, Boardman C. *Research Collaboration and Team Science: A State-Of-The-Art Review and Agenda*. Cham, Switzerland: Springer; 2014.
 17. Hall KL, Vogel AL, Huang GC, Serrano KJ, Rice EL, Tsakraklides SP, et al. The science of team science: a review of the empirical evidence and research gaps on collaboration in science. *Am Psychol* 2018;73(4):532-548. [doi: [10.1037/amp0000319](https://doi.org/10.1037/amp0000319)] [Medline: [29792466](https://pubmed.ncbi.nlm.nih.gov/29792466/)]
 18. Cooke NJ, Hilton ML. *Enhancing the Effectiveness of Team Science*. Washington, DC, USA: The National Academies Press; 2015.
 19. Hall KL, Stokols D, Moser RP, Taylor BK, Thornquist MD, Nebeling LC, et al. The collaboration readiness of transdisciplinary research teams and centers findings from the National Cancer Institute's TREC Year-One evaluation study. *Am J Prev Med* 2008 Aug;35(2 Suppl):S161-S172 [FREE Full text] [doi: [10.1016/j.amepre.2008.03.035](https://doi.org/10.1016/j.amepre.2008.03.035)] [Medline: [18619396](https://pubmed.ncbi.nlm.nih.gov/18619396/)]
 20. Mâsse LC, Moser RP, Stokols D, Taylor BK, Marcus SE, Morgan GD, et al. Measuring collaboration and transdisciplinary integration in team science. *Am J Prev Med* 2008 Aug;35(2 Suppl):S151-S160. [doi: [10.1016/j.amepre.2008.05.020](https://doi.org/10.1016/j.amepre.2008.05.020)] [Medline: [18619395](https://pubmed.ncbi.nlm.nih.gov/18619395/)]
 21. Oetzel JG, Zhou C, Duran B, Pearson C, Magarati M, Lucero J, et al. Establishing the psychometric properties of constructs in a community-based participatory research conceptual model. *Am J Health Promot* 2015;29(5):e188-e202 [FREE Full text] [doi: [10.4278/ajhp.130731-QUAN-398](https://doi.org/10.4278/ajhp.130731-QUAN-398)] [Medline: [24720389](https://pubmed.ncbi.nlm.nih.gov/24720389/)]
 22. Misra S, Stokols D, Cheng L. The transdisciplinary orientation scale: factor structure and relation to the integrative quality and scope of scientific publications. *J Transl Med Epidemiol* 2015;3(2):1042.
 23. Researcher survey form. Center for Clinical and Translational Science and Training. 2002. URL: https://cctst.uc.edu/sites/default/files/cis/survey-TTURC_research.pdf [accessed 2022-04-02]
 24. Stokols D, Fuqua J, Gress J, Harvey R, Phillips K, Baezconde-Garbanati L, et al. Evaluating transdisciplinary science. *Nicotine Tob Res* 2003 Dec;5 Suppl 1:S21-S39. [doi: [10.1080/14622200310001625555](https://doi.org/10.1080/14622200310001625555)] [Medline: [14668085](https://pubmed.ncbi.nlm.nih.gov/14668085/)]
 25. Baruch Y, Holtom BC. Survey response rate levels and trends in organizational research. *Hum Relat* 2008 Aug 01;61(8):1139-1160. [doi: [10.1177/0018726708094863](https://doi.org/10.1177/0018726708094863)]
 26. Kraska-Miller M. *Nonparametric Statistics for Social and Behavioral Sciences*. Boca Raton, FL, USA: Chapman and Hall/CRC; 2014.
 27. Mayring P. *Qualitative Content Analysis: Theoretical Foundation, Basic Procedures and Software Solution*. Klagenfurt, Austria: Open Access Repository; 2014.
 28. Lungeanu A, Huang Y, Contractor NS. Understanding the assembly of interdisciplinary teams and its impact on performance. *J Informetr* 2014 Jan;8(1):59-70 [FREE Full text] [doi: [10.1016/j.joi.2013.10.006](https://doi.org/10.1016/j.joi.2013.10.006)] [Medline: [24470806](https://pubmed.ncbi.nlm.nih.gov/24470806/)]
 29. Conway JM, Lance CE. What reviewers should expect from authors regarding common method bias in organizational research. *J Bus Psychol* 2010 May 22;25(3):325-334. [doi: [10.1007/s10869-010-9181-6](https://doi.org/10.1007/s10869-010-9181-6)]
 30. Jordan PJ, Troth AC. Common method bias in applied settings: the dilemma of researching in organizations. *Aust J Manag* 2019 Sep 03;45(1):3-14. [doi: [10.1177/0312896219871976](https://doi.org/10.1177/0312896219871976)]
 31. Chouvarda I, Mountford N, Trajkovik V, Loncar-Turukalo T, Cusack T. Leveraging interdisciplinary education toward securing the future of connected health research in Europe: qualitative study. *J Med Internet Res* 2019 Nov 13;21(11):e14020 [FREE Full text] [doi: [10.2196/14020](https://doi.org/10.2196/14020)] [Medline: [31719026](https://pubmed.ncbi.nlm.nih.gov/31719026/)]
 32. Vogel AL, Stipelman BA, Hall KL, Nebeling L, Stokols D, Spruijt-Metz D. Pioneering the transdisciplinary team science approach: lessons learned from National Cancer Institute grantees. *J Transl Med Epidemiol* 2014;2(2):1027 [FREE Full text] [Medline: [25554748](https://pubmed.ncbi.nlm.nih.gov/25554748/)]
 33. Koh D, Lee K, Joshi K. Transformational leadership and creativity: a meta-analytic review and identification of an integrated model. *J Organ Behav* 2019 Feb 08;40(6):625-650. [doi: [10.1002/job.2355](https://doi.org/10.1002/job.2355)]
 34. Wooten KC, Calhoun WJ, Bhavnani S, Rose RM, Ameredes B, Brasier AR. Evolution of multidisciplinary translational teams (MTTs): insights for accelerating translational innovations. *Clin Transl Sci* 2015 Oct;8(5):542-552 [FREE Full text] [doi: [10.1111/cts.12266](https://doi.org/10.1111/cts.12266)] [Medline: [25801998](https://pubmed.ncbi.nlm.nih.gov/25801998/)]

Abbreviations

ARAILIS: Augmented Reality and Artificial Intelligence Supported Laparoscopic Imagery in Surgery

OR: operating room

PROSPER: Platform for Operation Scheduling and Prediction Using Machine Learning

SRCC: Spearman rank correlation coefficient

TTS: transdisciplinary team science

Edited by A Kushniruk; submitted 19.01.22; peer-reviewed by Z Su; comments to author 09.02.22; revised version received 25.02.22; accepted 06.03.22; published 04.05.22.

Please cite as:

Krause-Jüttler G, Weitz J, Bork U

Interdisciplinary Collaborations in Digital Health Research: Mixed Methods Case Study

JMIR Hum Factors 2022;9(2):e36579

URL: <https://humanfactors.jmir.org/2022/2/e36579>

doi: [10.2196/36579](https://doi.org/10.2196/36579)

PMID: [35507400](https://pubmed.ncbi.nlm.nih.gov/35507400/)

©Grit Krause-Jüttler, Jürgen Weitz, Ulrich Bork. Originally published in JMIR Human Factors (<https://humanfactors.jmir.org>), 04.05.2022. This is an open-access article distributed under the terms of the Creative Commons Attribution License (<https://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Human Factors, is properly cited. The complete bibliographic information, a link to the original publication on <https://humanfactors.jmir.org>, as well as this copyright and license information must be included.

Original Paper

Proposal for Post Hoc Quality Control in Instrumented Motion Analysis Using Markerless Motion Capture: Development and Usability Study

Hanna Marie Röhling^{1,2,3,4}, MSc; Patrik Althoff^{1,2,3}; Radina Arsenova^{1,2,3,5}, MD; Daniel Drebinge^{1,2,3}; Norman Gigengack^{1,2,3}; Anna Chorschew^{1,2,3}; Daniel Kroneberg⁶, MD; Maria Rönnefarth^{6,7}, MD; Tobias Ellermeyer^{1,2,3,8}; Sina Cathérine Rosenkranz^{9,10}, MD; Christoph Heesen^{9,10}, MD; Behnoush Behnia¹¹, MD; Shigeki Hirano¹², MD, PhD; Satoshi Kuwabara¹², MD, PhD; Friedemann Paul^{1,2,3,6,13}, MD; Alexander Ulrich Brandt^{13,14}, MD; Tanja Schmitz-Hübsch^{1,2,3,13}, MD

¹Experimental and Clinical Research Center, a cooperation between the Max-Delbrück-Center for Molecular Medicine in the Helmholtz Association and the Charité - Universitätsmedizin Berlin, Berlin, Germany

²Experimental and Clinical Research Center, Charité – Universitätsmedizin Berlin, corporate member of Freie Universität Berlin and Humboldt Universität zu Berlin, Berlin, Germany

³Max-Delbrück-Center for Molecular Medicine in the Helmholtz Association (MDC), Berlin, Germany

⁴Motognosis GmbH, Berlin, Germany

⁵Department of Pediatrics, St Joseph Krankenhaus Berlin-Tempelhof, Berlin, Germany

⁶Department of Neurology, Charité – Universitätsmedizin Berlin, corporate member of Freie Universität Berlin and Humboldt Universität zu Berlin, Berlin, Germany

⁷Clinical Study Center, Berlin Institute of Health at Charité – Universitätsmedizin Berlin, Berlin, Germany

⁸Department of Neurology, Vivantes Auguste-Viktoria-Klinikum, Berlin, Germany

⁹Institute of Neuroimmunology and Multiple Sclerosis, Center for Molecular Neurobiology Hamburg (ZMNH), University Medical Center Hamburg-Eppendorf, Hamburg, Germany

¹⁰Department of Neurology, University Medical Center Hamburg-Eppendorf, Hamburg, Germany

¹¹Department of Psychiatry, Charité – Universitätsmedizin Berlin, corporate member of Freie Universität Berlin and Humboldt Universität zu Berlin, Berlin, Germany

¹²Department of Neurology, Graduate School of Medicine, Chiba University, Chiba, Japan

¹³NeuroCure Clinical Research Center, Charité – Universitätsmedizin Berlin, corporate member of Freie Universität Berlin and Humboldt Universität zu Berlin, Berlin, Germany

¹⁴Department of Neurology, University of California, Irvine, CA, United States

Corresponding Author:

Hanna Marie Röhling, MSc

Experimental and Clinical Research Center, a cooperation between the Max-Delbrück-Center for Molecular Medicine in the Helmholtz Association and the Charité - Universitätsmedizin Berlin

Lindenberger Weg 80

Berlin, 13125

Germany

Phone: 49 30 450539718

Fax: 49 30 450 539915

Email: hanna-marie.roehling@charite.de

Abstract

Background: Instrumented assessment of motor symptoms has emerged as a promising extension to the clinical assessment of several movement disorders. The use of mobile and inexpensive technologies such as some markerless motion capture technologies is especially promising for large-scale application but has not transitioned into clinical routine to date. A crucial step on this path is to implement standardized, clinically applicable tools that identify and control for quality concerns.

Objective: The main goal of this study comprises the development of a systematic quality control (QC) procedure for data collected with markerless motion capture technology and its experimental implementation to identify specific quality concerns and thereby rate the usability of recordings.

Methods: We developed a post hoc QC pipeline that was evaluated using a large set of short motor task recordings of healthy controls (2010 recordings from 162 subjects) and people with multiple sclerosis (2682 recordings from 187 subjects). For each of these recordings, 2 raters independently applied the pipeline. They provided overall usability decisions and identified technical and performance-related quality concerns, which yielded respective proportions of their occurrence as a main result.

Results: The approach developed here has proven user-friendly and applicable on a large scale. Raters' decisions on recording usability were concordant in 71.5%-92.3% of cases, depending on the motor task. Furthermore, 39.6%-85.1% of recordings were concordantly rated as being of satisfactory quality whereas in 5.0%-26.3%, both raters agreed to discard the recording.

Conclusions: We present a QC pipeline that seems feasible and useful for instant quality screening in the clinical setting. Results confirm the need of QC despite using standard test setups, testing protocols, and operator training for the employed system and by extension, for other task-based motor assessment technologies. Results of the QC process can be used to clean existing data sets, optimize quality assurance measures, as well as foster the development of automated QC approaches and therefore improve the overall reliability of kinematic data sets.

(*JMIR Hum Factors* 2022;9(2):e26825) doi:[10.2196/26825](https://doi.org/10.2196/26825)

KEYWORDS

instrumented motion analysis; markerless motion capture; visual perceptive computing; quality control; quality reporting; gait analysis

Introduction

With technology rapidly advancing, instrumented motion analysis (IMA) has emerged as an auspicious tool to augment clinical decision-making in persons with motor impairments [1-5]. Applications range from complex gait laboratory equipment to consumer grade health apps, which quantify what a person can do in a standardized setting (motor capacity) or what a person does in everyday life (motor performance) [6]. Regarding motor capacity, marker-based optoelectronic motion analysis systems serve as the gold standard for other technologies [7,8] and are, for instance, successfully used in treatment planning for children with cerebral palsy [9]. However, their high cost and complexity of analysis comprise significant disadvantages for clinical use. Thus, technologies that are portable, affordable, and easy to use are more promising for large-scale application. Respective devices developed for clinical use include pressure-sensitive walkways, inertial sensors ("wearables"), and markerless motion capture systems based on consumer depth cameras [2,10]. In the following, the term IMA will be used for this more versatile subcategory of motion analysis systems.

Despite favorable properties, IMA has not been successfully integrated into wide clinical routine yet [11,12]. Although regulatory requirements for medical products address safety and accuracy within the context of use (eg, for application in specific diseases) [13-15], successful implementation of IMA further depends on acceptance from patients and clinicians. Thus, technical usability, interpretability of outcomes, and quantifiable clinical benefits play a major role in this development. Standardized and efficient quality control (QC) procedures, not only during initial development but also during advancement and application of a system, could facilitate this technological maturation process. We found such QC aspects to be largely understudied and underreported.

QC can be applied at three levels: preventive, ad hoc, and post hoc. Preventive QC is applied before data acquisition. Manufacturers or developing groups generate initial results on data quality and publish them in proof-of-concept studies, including small samples of healthy subjects and target groups for clinical application [7,8,16,17]. Such studies can identify major pitfalls and elaborate on correct usage of these systems. For technology that is already in use with a substantial number of researchers or clinicians, expert consensus can further yield guidelines to improve preventive QC [18]. Ad hoc QC is pertained during measurements. Depending on the system, operators can decide to discard, reinstruct, and rerecord upon observing deviations from standard operating procedures (SOPs) or receiving error messages. Lastly, post hoc QC is employed at the data analysis stage. One option in this context is univariate or multivariate outlier analysis based on the kinematic parameters [19-21]. However, these approaches are highly data-dependent, inept to uncover systematic errors or "false normal" parameter values, and do not provide information regarding underlying causes of data deviation. Additional post hoc QC measures constitute postprocessing tools and successive recalculation of kinematic parameters [22,23] as well as plausibility checks based on raw data [24-26]. To date, such processes have only been performed on comparatively small data sets.

In this study, we used data acquired with the emerging Motognosis Labs system (Motognosis GmbH) that extracts kinematic parameters from depth camera recordings. In recent years, this system was extensively used in a research context at our site and our cooperating sites [24-29] with a standardized protocol for short motor tasks specifically designed to assess motor capacities of people with multiple sclerosis (MS) [7,30]. Regarding preventive QC, previously established SOPs for system operators and patient instructions were used for all data analyzed herein. With respect to ad hoc QC, the software provides visual feedback regarding general subject positioning

in the volume of acquisition and real-time tracking of the whole body as well as individual body parts. Regarding post hoc QC, we found previously employed approaches to be either insufficient, incomplete, or not feasible to reliably examine large amounts of data [19-21,24-26]. Likewise, review of IMA literature did not yield any standards or generalizable concepts. Thus, we propose an approach for systematic post hoc QC, enabling clinical users to prevent, detect, and eliminate data of inferior quality.

For the quality concerns considered here, we distinguish technical and performance issues. Technical issues comprise system-specific malfunctioning of hardware and software as well as artifacts specific to the recording technique, such as signal interference due to subjects' clothing or the recording environment in the case of depth sensing technology. Performance issues can be considered less technology-specific and can be attributed either to the operator (eg, by providing faulty instructions) or to noncompliance of the recorded subject. If the latter is unrelated to the disease, it should lead to trial exclusion; however, impairment-related inability can be considered a feature of interest.

The main objectives of this study were to (1) build a post hoc QC pipeline that is efficient, user-friendly, and adaptable, enabling clinical users to make standardized and robust decisions concerning usability of individual recordings; (2) perform QC for a large number of recordings acquired at different study sites and thus investigate the types and frequencies of quality issues; and (3) analyze the feasibility of the approach.

Methods

Data Set

Our study was based on recordings of short, structured motor tasks captured with the Motognosis Labs system. This system relies on a consumer depth camera (Microsoft KinectV2, Microsoft Corporation) and visual perceptive computing. More precisely, the software development kit associated with the camera allows for the markerless tracking of 3D time series from 25 artificial anatomical landmarks for subjects located at 1.5 to 4.5 m from the camera. Custom Motognosis Labs algorithms employ these time series to extract kinematic parameters to quantify various aspects of motor capacity.

Data were pooled from 8 monocentric studies at 3 study sites that used software versions 1.1, 1.4, 2.0, or 2.1 as part of their

protocols. These studies will be referred to using the following identifiers: ASD, CIS, Valkinect, VIMS, and WALKIMS-DA (conducted at Charité – Universitätsmedizin Berlin, Berlin, Germany); Ambos and Oprims (conducted at Universitätsklinikum Eppendorf, Hamburg, Germany); and Chiba (conducted at Chiba University, Chiba, Japan). These studies were approved by the respective institutional review boards and all subjects provided written informed consent. The data set comprised recordings from 187 persons with MS and 162 healthy controls. VIMS, Valkinect, and WALKIMS-DA included both groups, whereas the other studies contributed subjects from 1 group only. Descriptive statistics include information on gender, age, anthropometry, and disease severity in case of people with MS, as measured by the Expanded Disability Status Scale [31] (Table 1 and study-specific information in Table S1 in Multimedia Appendix 1).

All subjects performed the Perceptive Assessment in Multiple Sclerosis (PASS-MS) protocol or parts of it between December 2014 and April 2019. PASS-MS consists of 10 structured motor tasks: Postural Control (POCO), Postural Control with Dual Task (POCO-DUAL), Stepping in Place (SIP), Stand Up and Sit Down (SAS), Short Line Walk (SLW), Short Comfortable Speed Walk (SCSW), Short Maximum Speed Walk (SMSW), Pronator Drift Test, Finger-Nose Test, and Finger Tapping. The latter 3 tasks were excluded from this study, as evaluation algorithms were still in an explorative stage at the time, yielding premature claims regarding data quality. A description of the remaining tasks except POCO-DUAL can be found in Otte et al [7,30]. POCO-DUAL equates to POCO with the addition of a cognitive task (Serial 3's subtraction). System operators had received in-depth training on how to use Motognosis Labs according to written SOPs. System SOPs included specifications of the setup, subject instructions, and rejection guidelines for recordings affected by performance and technical issues. According to the protocol, SAS, SLW, SCSW, and SMSW are recorded thrice consecutively, whereas POCO, POCO-DUAL, and SIP are recorded once. Deviations from SOPs occurred when single tasks or task repetitions were omitted, or operators decided to produce additional recordings (all of which should prompt an operator comment that is stored along with raw data of each recording). Such deviations explain incongruencies in the numbers of recordings per task (Table 1 and study-specific information in Table S2 in Multimedia Appendix 1), as all available recordings were included in this post hoc QC initiative.

Table 1. Demographic information about study subjects with missing data indicated as percentages and number of recordings per Perceptive Assessment in Multiple Sclerosis task subdivided by disease status.

Subject characteristics	All	HC ^a	PwMS ^b
Demographics			
N (% female; % — ^c)	349 (51.6; 0.6)	162 (51.2; 1.2)	187 (51.9; 0)
Age (years), mean (SD; % —)	42.0 (12.2; 0.6)	38.3 (12.8; 1.2)	45.3 (10.8; 0)
Height (cm), mean (SD; % —)	173.1 (9.2; 2.6)	172.0 (9.6; 3.7)	174.1 (8.8; 1.6)
Weight (kg), mean (SD; % —)	72.9 (14.8; 8.0)	70.4 (14.6; 8.0)	75.0 (14.6; 8.0)
BMI (kg/m ²), mean (SD; % —)	24.3 (4.1; 8.0)	23.8 (3.9; 8.0)	24.7 (4.3; 8.0)
EDSS ^d median (range; % —)	N/A ^e	N/A	3.0 (0.0-6.5; 2.7)
# of recordings per PASS-MS^f task			
All	4692	2010	2682
POCO ^g	354	165	189
POCO-DUAL ^h	245	88	157
SCSW ⁱ	1043	489	554
SMSW ^j	907	361	546
SLW ^k	957	428	529
SIP ^l	291	131	160
SAS ^m	895	348	547

^aHC: healthy controls.^bPwMS: people with multiple sclerosis.^c—: not available.^dEDSS: Expanded Disability Status Scale.^eN/A: not applicable.^fPASS-MS: Perceptive Assessment in Multiple Sclerosis.^gPOCO: Postural Control.^hPOCO-DUAL: Postural Control with Dual Task.ⁱSCSW: Short Comfortable Speed Walk.^jSMSW: Short Maximum Speed Walk.^kSLW: Short Line Walk.^lSIP: Stepping in Place.^mSAS: Stand Up and Sit Down.

QC Pipeline Development

The QC pipeline development comprised 2 key components. First, we implemented informative visualizations enabling raters to classify the quality of raw data from PASS-MS recordings and hence implicitly assess the reliability of associated kinematic parameters. Second, we developed an efficient rating strategy for large numbers of recordings.

For the creation of informative visualizations, videos from raw depth streams were generated to enable review of each recorded task. The depth information was further used to produce a condensed representation of each recording in the form of 3 images that are hereafter referred to as motion profiles. They comprise images of depth data averaged over time, over the vertical direction, and over the horizontal direction. As PASS-MS tasks are short and highly standardized, we assumed

that major protocol deviations and technical issues would be easily identifiable from motion profiles. To allow for the detection of more subtle quality issues, we also illustrated characteristic signals that are used to calculate kinematic parameters with Motognosis Labs. Visualizations were generated using Python (version 3.7.3) and the matplotlib package (version 3.1.0). A stratified random sample from 15 people with MS and 14 healthy controls was used to test and update visualizations and determine the main rating criteria per task.

We then built a graphical user interface (GUI), which includes a rating window containing visualizations, an overall usability decision checkbox (keep, discard, undecided), and task-specific multiselect checkboxes containing the main rating criteria. Furthermore, on-demand viewers for depth videos and operator comments were integrated. The GUI was programmed in Python (version 3.7.3) using the tkinter package (version 8.6). We

prepared detailed rating manuals as well as oral instructions (~45 minutes) to familiarize raters with the GUI. The entire data set (see [Table 1](#)) was subjected to ratings, such that each recording was investigated by 2 independent raters. In this step, 8 raters evaluated a total of 4692 recordings from 162 healthy controls and 187 people with MS. Raters comprised medical students, clinician scientists or researchers in other professions, and trained neurologists, all from Charité, Berlin. Among them, 6 raters had operated Motognosis Labs before, whereas 2 were new to the system. Moreover, 2 raters had been actively involved in the development of the QC pipeline, whereas 6 were new to any systematic QC of the data. After in-depth instructions, ratings were conducted individually by the raters at a self-selected speed.

Statistical Analysis

Statistical analyses included the extraction of frequencies for overall usability decisions, rater concordance and discordance, and selected rating criteria. The former 2 were illustrated as

confusion matrices. Furthermore, the median rating duration per recording was extracted from the GUI log files. Figures were produced with Python (version 3.7.3) using the matplotlib package (version 3.1.0).

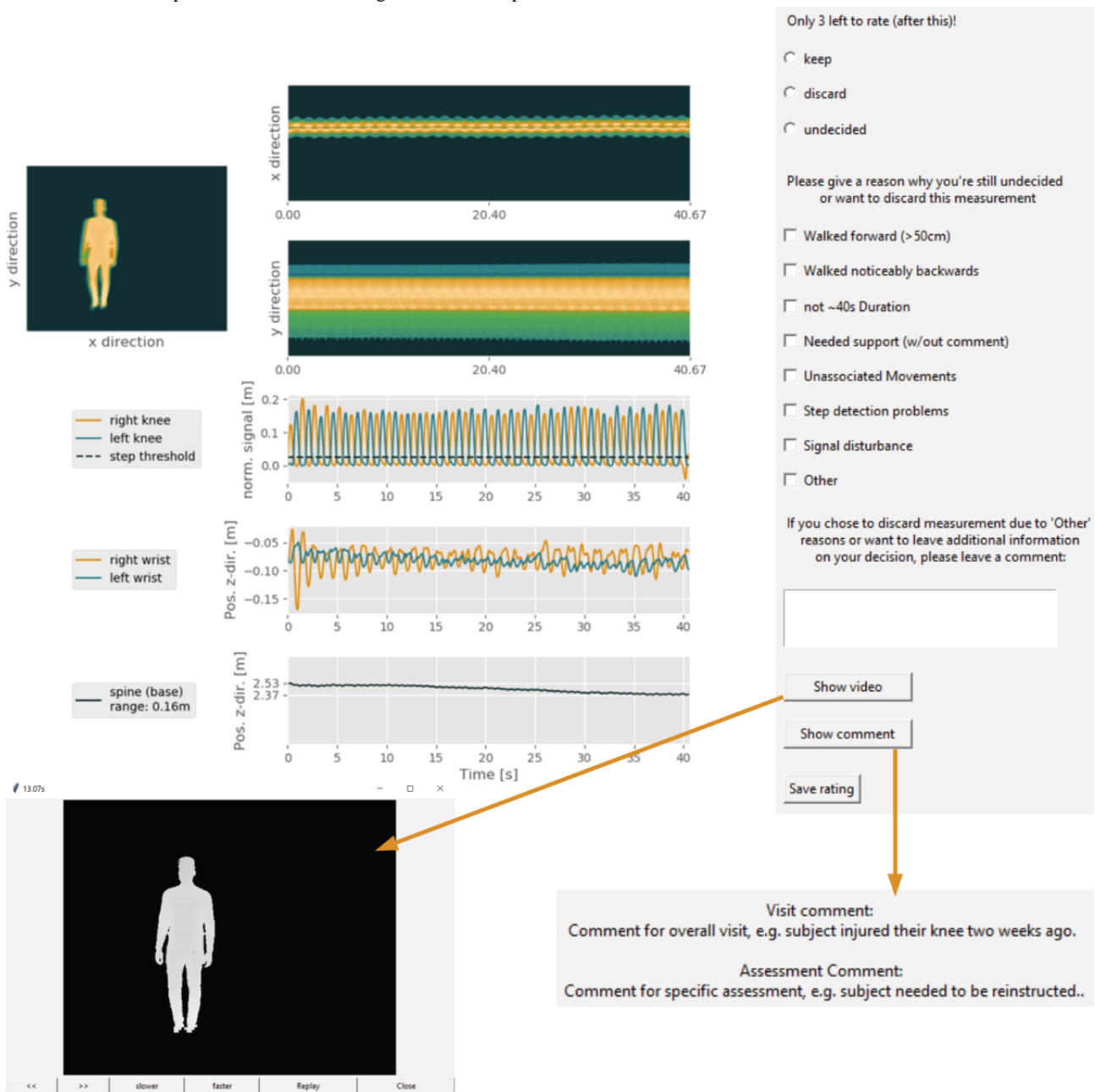
Results

QC Pipeline Usage and Feasibility

After generating visualizations, the implemented GUI can be opened to progressively rate motor task recordings. Intermediate results can be saved in an underlying Excel file, such that raters can flexibly organize their workload. An example of the rating window including respective visualizations, checkboxes, and buttons is shown in [Figure 1](#).

Oral feedback from raters upon completion confirmed that the GUI and the QC pipeline behind it were easy to use and effective. The median rating duration per recording amounted to 6.3 seconds.

Figure 1. Rating window screenshots for an exemplary Stepping in Place recording. Upper left: motion profiles generated by summation of frontally recorded depth data over time, along horizontal and vertical directions and signal curves characteristic of the task (here: knee amplitudes, arm sway, and overall subject positioning over time). Upper right: checkboxes for usability decisions and main criteria including an option for free-text comments. Lower left: on-demand depth video viewer. Lower right: on-demand operator comment viewer.

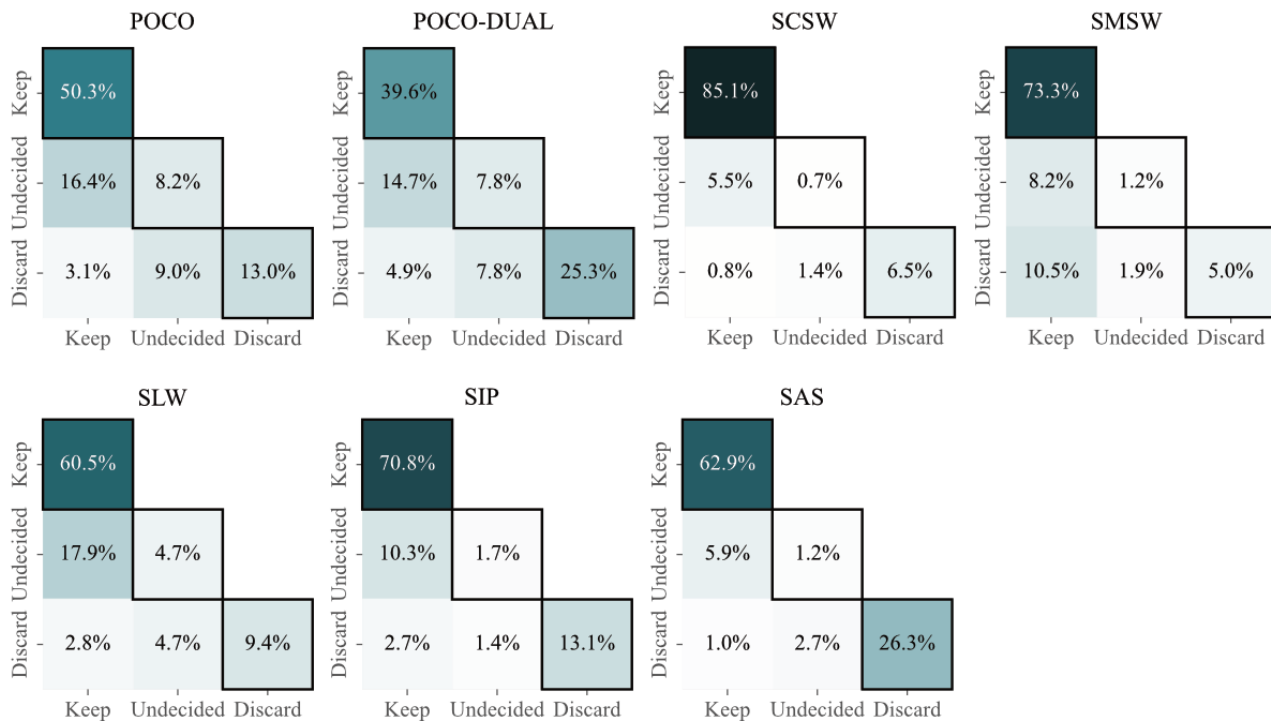


Rater Concordance and Usability of Recordings

Concerning keep, discard, or undecided decisions, raters concurred on more than 70% of recordings for each task (POCO: 71.5%, POCO-DUAL: 72.7%, SCSW: 92.3%, SMSW: 79.5%, SLW: 74.6%, SIP: 85.6%, and SAS: 90.4%) (Figure 2). Consequently, we observed discordance for up to 28.5% of

recordings, which points to task-specific difficulties in using the rating criteria. However, such discordance was mostly due to 1 rater’s undecided decision. Instances of strictly opposing usability, meaning that 1 rater voted keep and the other discard, were uncommon (between 0.8% and 4.9%), except for SMSW (10.5%).

Figure 2. Synopsis of usability decisions by 2 raters per recording per Perceptive Assessment in Multiple Sclerosis task. Rater agreement on usability decisions keep, discard, and undecided are framed. POCO: Postural Control; POCO-DUAL: Postural Control with Dual Task; SAS: Stand Up and Sit Down; SCSW: Short Comfortable Speed Walk; SIP: Stepping in Place; SLW: short line walk; SMSW: Short Maximum Speed Walk.



A task-wise visualization of rater decisions regarding usability of recordings is depicted in Figure 2. Unobjectionable usability, defined as a unanimous keep decision, was obtained for 85.1% of SCSW, more than 70% of SMSW and SIP (73.3% and 70.8%, respectively), more than 60% for SAS and SLW (62.9% and 60.5%, respectively) and less than or close to half for POCO and POCO-DUAL recordings (50.3% and 39.6%, respectively). The highest rates for unanimous discard decisions were observed for SAS (26.3%), followed by POCO-DUAL (25.3%), and POCO and SIP (13.0% and 13.1%, respectively). The respective rates were low for gait tasks including SLW, SCSW, and SMSW (9.4%, 6.5%, and 5.0%, respectively). Rater concordance as well as proportions of unanimous keep and discard decisions subdivided for all studies can be found in Table S3 in Multimedia Appendix 1.

Main Quality Concerns

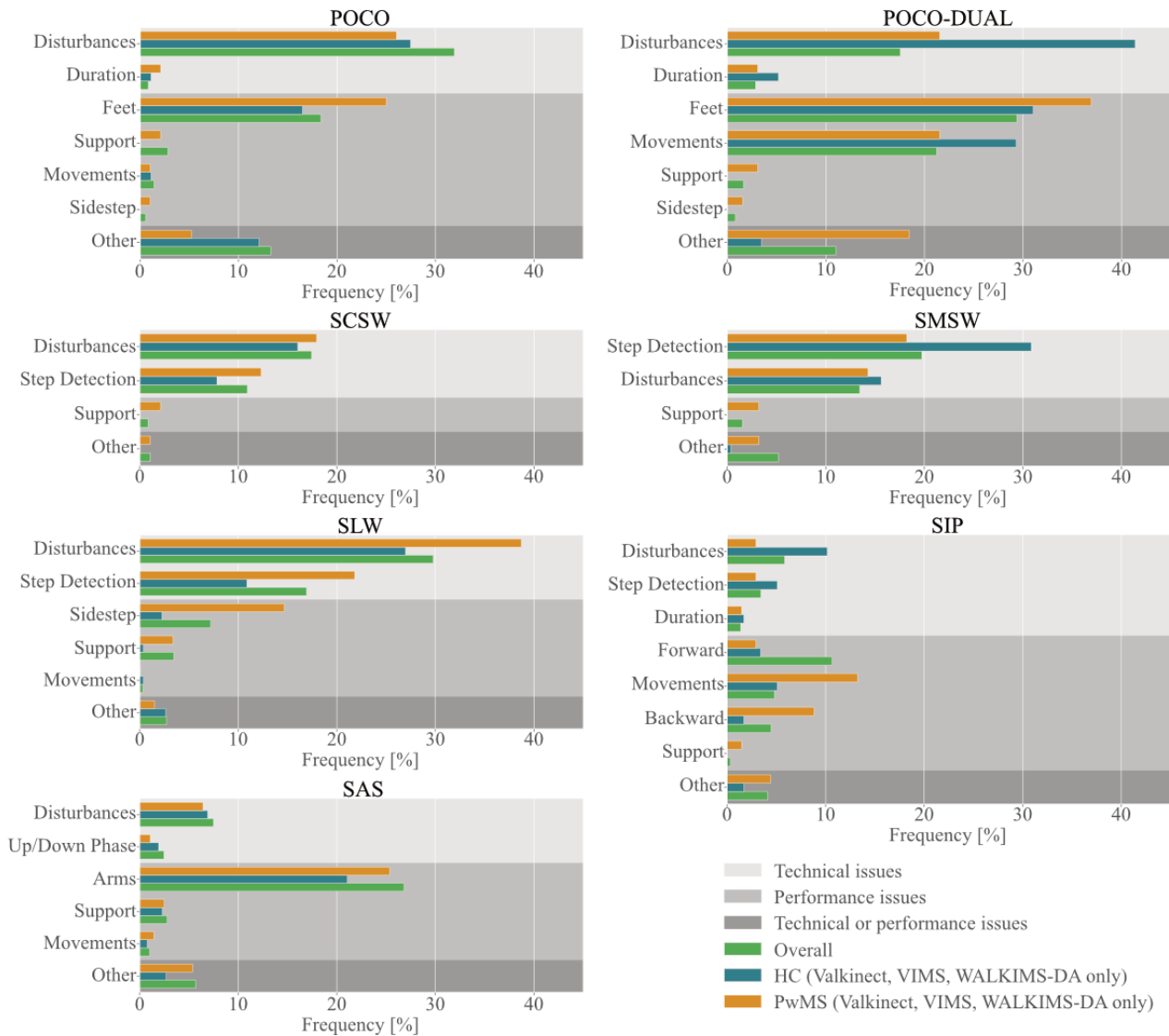
The main rating criteria compiled during QC pipeline development are listed below, with the respective tasks indicated in parentheses.

- Disturbances, technical issue: Signal disturbances including noisy background, floor, and technical issues with tracking clothing (all tasks)
- Duration, technical issue: Recording duration substantially deviating from 40 seconds, namely a deviation of more than 1 second (POCO, POCO-DUAL, and SIP)
- Step Detection, technical issue: Incorrect Step Detection (SCSW, SMSW, SIP, and SLW)

- Up/Down Phase, technical issue: Incomplete or incorrectly detected standing-up or sitting-down phase (SAS)
- Arms, performance issue: Arms not hanging loosely down at the beginning of the recording (SAS)
- Backward, performance issue: Subject walking backward by more than 50 cm or exhibiting a deliberate backward correction (SIP)
- Feet, performance issue: Deviation from closed feet position, namely if the feet are in an open or a V-shaped position (POCO and POCO-DUAL)
- Forward, performance issue: Subject moving forward by more than 50 cm (SIP)
- Movements, performance issue: Task-unassociated movements such as scratching or gesturing (POCO, POCO-DUAL, SLW, SIP, and SAS)
- Sidestep, performance issue: 1 or multiple sidesteps (POCO, POCO-DUAL, and SLW)
- Support, performance issue: Subject needing support from a walking stick, walls, rollator, or the like (all tasks)
- Other, technical or performance issue: Other/unlisted criterion (all tasks)

Respective selection frequencies (multiple selections were possible) are illustrated in Figure 3. Possible disease-associated differences in data quality can be estimated from the 3 studies featuring healthy controls and people with MS, namely VIMS, Valkinect, and WALKIMS-DA.

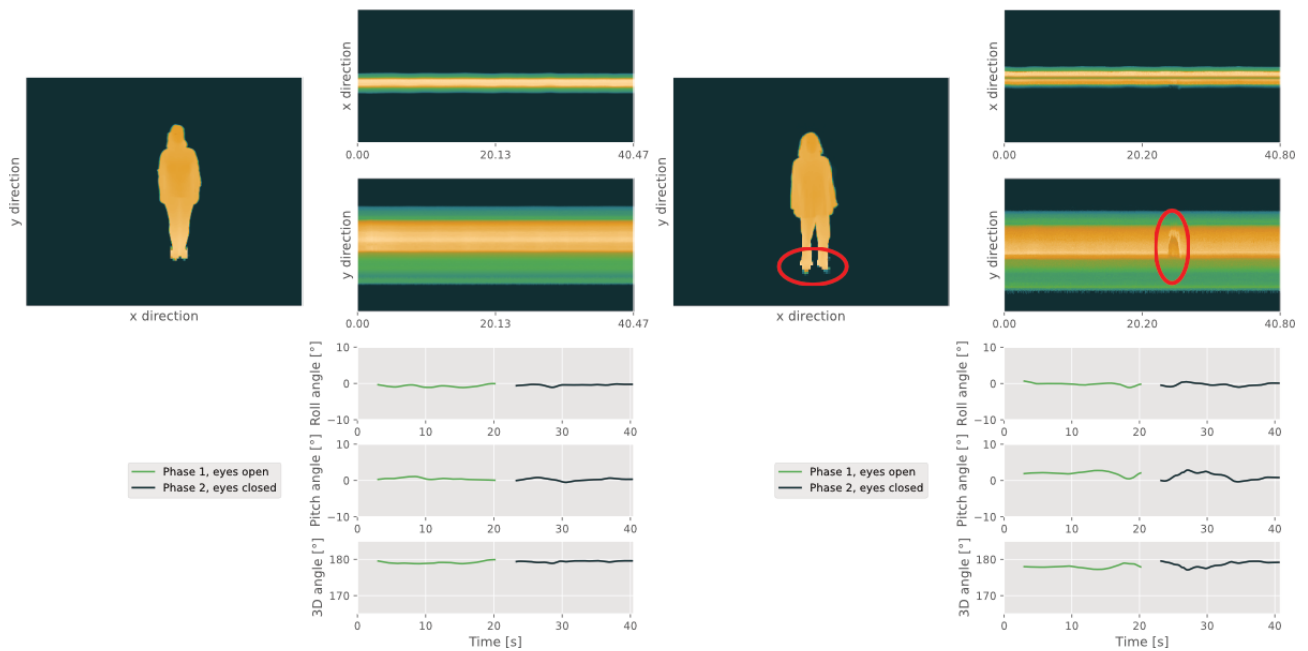
Figure 3. Selection frequencies of technical and performance-related rating criteria for all subjects as well as split by group for the 3 studies featuring healthy controls and people with multiple sclerosis. HC: healthy controls; POCO: Postural Control; POCO-DUAL: Postural Control with Dual Task; PwMS: people with multiple sclerosis; SAS: Stand Up and Sit Down; SCSW: Short Comfortable Speed Walk; SIP: Stepping in Place; SLW: Short Line Walk; SMSW: Short Maximum Speed Walk.



The most prevalent quality concerns comprised Feet, Disturbances, and Other for POCO and additionally Movements for POCO-DUAL. An example of a POCO recording that was discarded due to incorrect Feet positioning as well as unassociated Movements, namely the most frequent performance-associated quality concerns, can be found in Figure 4. For POCO-DUAL, supposedly task-unassociated movements were tagged with Movements and Other by the raters. However,

these hand and arm movements often seemed to result from cognitive efforts made during mental arithmetic. In this case, no clear distinction between task-associated and task-unassociated movements can be made. Regarding technical quality concerns, raters' comments suggested that recordings tagged with Disturbances or Other most often exhibited noisy or corrupt leg, feet, or floor signals.

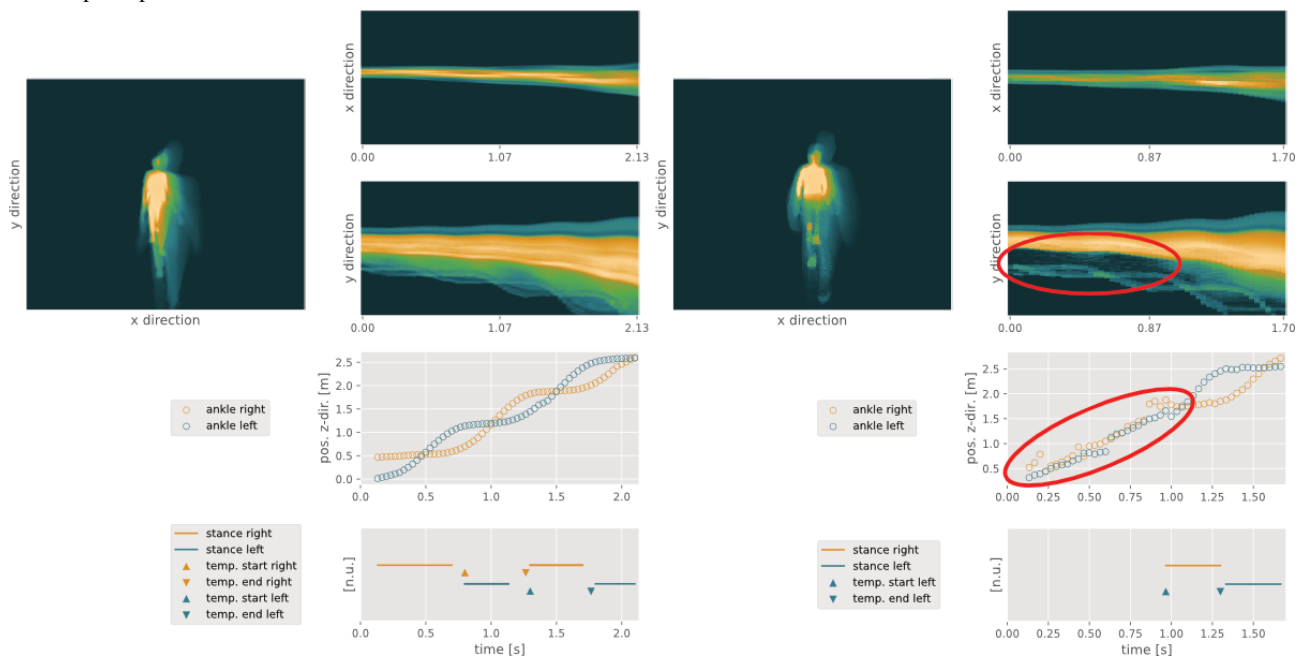
Figure 4. Left: quality control pipeline visualization screenshot of a high-quality Postural Control recording. Right: quality control pipeline visualization screenshot of a Postural Control recording featuring 2 frequently observed performance-related quality concerns, incorrect Feet positioning (according to standard operating procedures, the forefoot and heel should be closed) and unassociated hand Movements around second 22.



Prevalent quality concerns for gait tasks were Disturbances and Step Detection in SLW and— less frequently—SCSW and SMSW. A cross-dependency between the 2 criteria was often observed when unsuitable clothing led to noisy signals (noted as Disturbances by the raters), which in turn leads to issues

concerning Step Detection. An example of this issue for an SCSW recording is depicted in Figure 5. Other Disturbances related to floor reflections were not associated with Step Detection issues as often.

Figure 5. Left: quality control pipeline visualization screenshot of a high-quality short comfortable speed walk recording. Right: quality control pipeline visualization screenshot of a Short Comfortable Speed Walk recording featuring a frequently observed technical quality concern, unsuitable clothing causing Disturbances and thus Step Detection issues. Abbreviation temp. represents temporal and indicates the detected stance phases used for temporal rather than spatial parameters.



Excessive forward locomotion (Forward) was the most frequent quality concern for SIP recordings. However, from our experience, the chosen threshold of 50 cm forward motion is rather conservative and distances up to 80-100 cm might be tolerable.

The most prominent problem for SAS was incorrect arm positioning (Arms) at the beginning of a recording. Such incorrect arm positioning was not easily discernible from the motion profile alone and raters usually consulted the provided depth videos to confirm this specific quality concern. Furthermore, a mistake in signal plot generation for

SAS—affecting 3.8% of SAS plots—led to an overestimation of recordings affected by the Up/Down Phase criterion. [Figure 3](#) provides raw ratings, and the represented numbers hence reflect this overestimation.

Disparities between people with MS and healthy controls for performance-related quality aspects were apparent for the generally less often observed Support (all tasks) and Sidestep (POCO, POCO-DUAL, and SLW) issues. This can be interpreted as a disease-related difficulty or the inability to follow task instructions. Results regarding incorrect Feet positioning during POCO and POCO-DUAL did not allow for the interpretation of this criterion as a mainly disease-related one. This criterion as well as Forward and Backward motion during SIP and the incorrect starting position of the Arms during SAS were present in both groups, though slightly more frequent in people with MS. Frequencies of the observed quality criteria further subdivided for all studies can be found in [Table S4 in Multimedia Appendix 1](#).

Discussion

This study presents a post hoc QC pipeline for clinical users of an IMA system. Its core consists of an interface, which enables an intuitive usability decision for individual recordings based on an extendable set of quality criteria. The pipeline proved highly feasible for users—including raters less acquainted with the IMA system itself—and yielded acceptable rater concordance. Its application in a large set of recordings from healthy controls and people with MS demonstrated the utility and necessity of post hoc QC to ensure reliable data and avoid misinterpretation of IMA results. It further identified points for improvement in preventive and ad hoc QC. To our knowledge, this is the first study to systematically investigate QC aspects and propose a clinically applicable QC pipeline for visual perceptive computing.

In the following, we will discuss 2 main aspects of our results. First, the rater concordance, which indicates the feasibility and limitations of our QC approach, and second, the usability decisions themselves, which indicate the quality and limitations of our data.

Rater concordance between 71.5% to 92.3% was generally acceptable. Only for SMSW, strictly opposed keep/discard decisions occurred to a relevant extent (10.5%). This was mostly caused by 1 rater's discard decisions because no full gait cycle was captured. Due to the limited recording range of the depth camera, this is a frequent observation for SMSW and cannot be directly attributed to technical or performance issues. Generally, discordance may reflect ambiguity regarding rating criteria, difficulties in the evaluation of individual cases, or rater oversight. Probably only 1 rater, most likely the operator of the system, will apply post hoc QC in future clinical applications. Thus, possible reasons for rater discordance should be carefully addressed in further development of the QC pipeline, for instance, by specifying the rating criteria, as well as conducting more targeted rater trainings. However, as with other clinical judgments, QC decisions will remain informed, but ultimately intuitive decisions.

Usability decisions were interpreted as follows. Recordings receiving a unanimous keep or discard decision from the corresponding 2 raters were regarded as having assessable and satisfactory or unsatisfactory quality, respectively. Remaining recordings with discordant or undecided usability decisions were classified as needing further investigation, thus being less assessable and with potentially objectionable quality. The proportion of unanimous keep decisions varied substantially between tasks (39.6%-85.1%). In this respect, the SCSW task had the most favorable results with the highest rater concordance (92.3%) and the highest proportion of keep decisions among all tasks. At the other end of the spectrum were POCO and POCO-DUAL with rather moderate rater concordance (71.5% and 72.7%, respectively) and comparatively less unanimous keep decisions (50.3% and 39.6%, respectively). This partial ambiguity supports our inclusion of undecided as an option to avoid forced decisions as well as free text comments to enable marking of unexpected quality concerns.

Regarding technical quality issues, the short walk tasks SCSW, SMSW, and SLW suffered the most from unfavorable properties of clothing that hampered infrared light reflection [32]. POCO and POCO-DUAL often exhibited noisy and cutoff feet signals, attributable to a limited differentiation of feet and ground leading to unstable landmark estimations, as reported earlier [7]. Countermeasures include general recommendations toward subjects' clothing and flooring at the measurement site.

We expected performance-related quality concerns to be associated with physical limitations and thus the disease status to some extent. This seemed to apply to rating criteria Sidestep and Support. However, the more commonly observed performance-related issues (eg, Feet and Movements for POCO and POCO-DUAL, Arms for SAS, and Forward for SIP) occurred in healthy subjects as well. This implies that mistakes in task instruction or ad hoc QC occurred to a relevant degree, despite detailed SOPs and operator training. Even higher proportions of performance-related issues may be expected with wider clinical use or in unsupervised telemedical applications. Thus, further IMA development should aim to implement technical measures for automated real-time detection of performance issues and respective response plans (eg, reinstruction and repetition). Performance-related quality concerns may specifically apply to the assessment of motor capacity in a lab setting or in task-based assessments as opposed to the recently proposed IMA systems for continuous assessment of motor performance [4,5,15].

In the literature, we found generally sparse reporting of QC aspects for IMA. This includes reporting of unobjectionable data quality, which we assume to be unlikely. As an indicator of technical IMA system performance, some authors reported exclusion of IMA recordings due to seemingly blatant technical failures, with rates ranging from a few corrupted examples to recordings of 48.8% of the participants [21,22,33,34]. Unfortunately, respective proportions could not be provided for our data set, as we did not track recordings discarded ad hoc. Regarding data exclusion in postprocessing, outlier detection was the most frequent approach. For univariate outlier detection on normative gait and balance parameters in children, exclusion rates of 2.5% and 6% were reported [20,35]. A multivariate

outlier detection approach on kinematic gait data with successive expert evaluation identified erroneous Step Detection in 3.4% of the subjects [21], whereas a custom post hoc QC procedure applied on SMSW data obtained using Motognosis Labs led to exclusion of 6.7% of the recordings [24]. We consider the QC approach presented here to be rather conservative when compared to outlier detection. It is highly possible that significant quality concerns identified at the raw data level would not be detected by outlier analysis at the kinematic parameter level. For example, failure to stand with closed feet during POCO most likely results in reduced postural sway, which would be mistaken for higher postural stability in the respective subject at the kinematic outcome level.

Lastly, reporting of manual postprocessing, for example, using the GAITRite footfall labeling tool, is often limited to whether it was employed at all [22,36], and respective proportions are only seldom addressed [37].

Beyond IMA, the need for QC has been recognized for other technical procedures. In the context of MS research, magnetic resonance imaging and optical coherence tomography serve as examples for which recommendations have been made regarding standardized protocols, QC, and harmonious reporting thereof [38-42]. Therefore, we propose standardized reporting of IMA results to include information regarding the following: (1) number of recording failures during data acquisition; (2) type and amount of applied postprocessing, both technical and manual; (3) fraction of recordings undergoing QC; (4) fraction of recordings ultimately excluded from analysis (mention of respective causes would be highly valuable for future users)

Limitations of this study may include the decision to have each recording viewed by 2 out of 8 available raters; this limits formal interrater reliability analyses and does not assess individual

rater bias. However, we did not aim to establish interrater reliability but focused on obtaining generalizable estimates of rater concordance and determining the feasibility of the approach with a reasonably diverse set of raters. Further, other possible factors influencing usability of the recordings were not specifically analyzed. These include effect of the study site, population, system operators, as well as subjects' age, height, and weight. However, we consider QC results generalizable to and representative of routine applications because of the large size and heterogeneity of our sample. Differences in hardware were not tracked in this study (Kinect 2 sensors and laptops). Likewise, differences in software versions were disregarded because they were considered not substantial. However, recommendations regarding hardware and software may prospectively play a role in preventive QC in large-scale applications.

Regarding transferability, the visualizations employed here were specific to Motognosis Labs. However, appropriate visualizations have been implemented for other IMA systems as well. Examples include footprint depictions from pressure-sensitive walkways or acceleration illustrations from inertial sensors. Thus, we expect the general QC approach presented in this study to be transferable to other IMA systems. As for the observed quality concerns, technical issues are mostly or partially transferable to other depth camera- or visual sensor-based systems, respectively. The performance issues observed here are even more generalizable and thus highly informative for all researchers and clinicians using lab- or task-based IMA. The results of this study clearly support the need for QC of IMA data to ensure objectivity and enhance acceptance by clinical users and regulators alike. As a first step, this approach can advance consensus on the QC standards of different IMA systems and ultimately improve data quality.

Acknowledgments

We thank NeuroCure Clinical Research Center (NCRC), funded by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) under Germany's Excellence Strategy – EXC-2049 – 390688087 and Charité-BIH Clinical Study Center. Parts of this work were funded by a restricted research grant from Roche. DK is a participant in the BIH-Charité Clinician Scientist Program funded by the Charité –Universitätsmedizin Berlin and the Berlin Institute of Health. SCR was supported by the Clinician-Scientist Fellowship from the Stifterverband für die Deutsche Wissenschaft and the Hertie Network of Excellence in Clinical Neuroscience of the Gemeinnützige Hertie-Stiftung.

Conflicts of Interest

AUB is a shareholder and HMR is a paid part-time employee at Motognosis GmbH. The other authors declare no conflict of interest.

Multimedia Appendix 1

Further information on data set, usability decisions, and rating criteria statistics.

[[DOCX File , 110 KB](#) - [humanfactors_v9i2e26825_app1.docx](#)]

References

1. Petraglia F, Scarcella L, Pedrazzi G, Brancato L, Puers R, Costantino C. Inertial sensors versus standard systems in gait analysis: a systematic review and meta-analysis. *Eur J Phys Rehabil Med* 2019 Apr;55(2):265-280 [[FREE Full text](#)] [doi: [10.23736/S1973-9087.18.05306-6](https://doi.org/10.23736/S1973-9087.18.05306-6)] [Medline: [30311493](#)]

2. Muro-de-la-Herran A, Garcia-Zapirain B, Mendez-Zorrilla A. Gait analysis methods: an overview of wearable and non-wearable systems, highlighting clinical applications. *Sensors (Basel)* 2014 Feb;14(2):3362-3394 [FREE Full text] [doi: [10.3390/s140203362](https://doi.org/10.3390/s140203362)] [Medline: [24556672](https://pubmed.ncbi.nlm.nih.gov/24556672/)]
3. Pradhan C, Wuehr M, Akrami F, Neuhaeuser M, Huth S, Brandt T, et al. Automated classification of neurological disorders of gait using spatio-temporal gait parameters. *J Electromyogr Kinesiol* 2015 Apr;25(2):413-422. [doi: [10.1016/j.jelekin.2015.01.004](https://doi.org/10.1016/j.jelekin.2015.01.004)] [Medline: [25725811](https://pubmed.ncbi.nlm.nih.gov/25725811/)]
4. Frechette ML, Meyer BM, Tulipani LJ, Gurchiek RD, McGinnis RS, Sosnoff JJ. Next steps in wearable technology and community ambulation in multiple sclerosis. *Curr Neurol Neurosci Rep* 2019 Sep;19:80. [doi: [10.1007/s11910-019-0997-9](https://doi.org/10.1007/s11910-019-0997-9)] [Medline: [31485896](https://pubmed.ncbi.nlm.nih.gov/31485896/)]
5. Alexander S, Peryer G, Gray E, Barkhof F, Chataway J. Wearable technologies to measure clinical outcomes in multiple sclerosis: a scoping review. *Mult Scler* 2021 Oct;27(11):1643-1656 [FREE Full text] [doi: [10.1177/1352458520946005](https://doi.org/10.1177/1352458520946005)] [Medline: [32749928](https://pubmed.ncbi.nlm.nih.gov/32749928/)]
6. Maetzler W, Rochester L, Bhidayasiri R, Espay AJ, Sánchez-Ferro A, van Uem JMT. Modernizing daily function assessment in Parkinson's disease using capacity, perception, and performance measures. *Mov Disord* 2021 Jan;36(1):76-82. [doi: [10.1002/mds.28377](https://doi.org/10.1002/mds.28377)] [Medline: [33191498](https://pubmed.ncbi.nlm.nih.gov/33191498/)]
7. Otte K, Kayser B, Mansow-Model S, Verrel J, Paul F, Brandt AU, et al. Accuracy and reliability of the kinect version 2 for clinical measurement of motor function. *PLoS One* 2016;11(11):e0166532 [FREE Full text] [doi: [10.1371/journal.pone.0166532](https://doi.org/10.1371/journal.pone.0166532)] [Medline: [27861541](https://pubmed.ncbi.nlm.nih.gov/27861541/)]
8. Webster KE, Wittwer JE, Feller JA. Validity of the GAITRite® walkway system for the measurement of averaged and individual step parameters of gait. *Gait Posture* 2005 Dec;22(4):317-321. [doi: [10.1016/j.gaitpost.2004.10.005](https://doi.org/10.1016/j.gaitpost.2004.10.005)] [Medline: [16274913](https://pubmed.ncbi.nlm.nih.gov/16274913/)]
9. Wren TAL, Tucker CA, Rethlefsen SA, Gorton GEIII, Öunpuu S. Clinical efficacy of instrumented gait analysis: systematic review 2020 update. *Gait Posture* 2020 Jul;80:274-279. [doi: [10.1016/j.gaitpost.2020.05.031](https://doi.org/10.1016/j.gaitpost.2020.05.031)] [Medline: [32563727](https://pubmed.ncbi.nlm.nih.gov/32563727/)]
10. Morrison C, D'Souza M, Huckvale K, Dorn JF, Burggraaff J, Kamm CP, et al. Usability and acceptability of ASSESS MS: assessment of motor dysfunction in multiple sclerosis using depth-sensing computer vision. *JMIR Hum Factors* 2015 Jan;2(1):e11 [FREE Full text] [doi: [10.2196/humanfactors.4129](https://doi.org/10.2196/humanfactors.4129)] [Medline: [27025782](https://pubmed.ncbi.nlm.nih.gov/27025782/)]
11. Espay AJ, Hausdorff JM, Sánchez-Ferro Á, Klucken J, Merola A, Bonato P, Movement Disorder Society Task Force on Technology. A roadmap for implementation of patient-centered digital outcome measures in Parkinson's disease obtained using mobile health technologies. *Mov Disord* 2019 May;34(5):657-663 [FREE Full text] [doi: [10.1002/mds.27671](https://doi.org/10.1002/mds.27671)] [Medline: [30901495](https://pubmed.ncbi.nlm.nih.gov/30901495/)]
12. Brichetto G, Pedullà L, Podda J, Tacchino A. Beyond center-based testing: understanding and improving functioning with wearable technology in MS. *Mult Scler* 2019 Sep;25(10):1402-1411. [doi: [10.1177/1352458519857075](https://doi.org/10.1177/1352458519857075)] [Medline: [31502913](https://pubmed.ncbi.nlm.nih.gov/31502913/)]
13. Maetzler W, Klucken J, Horne M. A clinical view on the development of technology-based tools in managing Parkinson's disease. *Mov Disord* 2016 Sep;31(9):1263-1271. [doi: [10.1002/mds.26673](https://doi.org/10.1002/mds.26673)] [Medline: [27273651](https://pubmed.ncbi.nlm.nih.gov/27273651/)]
14. Stamate C, Magoulas GD, Kueppers S, Nomikou E, Daskalopoulos I, Jha A, et al. The cloudUPDRS app: a medical device for the clinical assessment of Parkinson's Disease. *Pervasive Mob Comput* 2018 Jan;43:146-166. [doi: [10.1016/j.pmcj.2017.12.005](https://doi.org/10.1016/j.pmcj.2017.12.005)]
15. Viceconti M, Hernandez Penna S, Dartee W, Mazzà C, Caulfield B, Becker C, et al. Toward a regulatory qualification of real-world mobility performance biomarkers in Parkinson's patients using digital mobility outcomes. *Sensors (Basel)* 2020 Oct;20(20):5920 [FREE Full text] [doi: [10.3390/s20205920](https://doi.org/10.3390/s20205920)] [Medline: [33092143](https://pubmed.ncbi.nlm.nih.gov/33092143/)]
16. Steinert A, Sattler I, Otte K, Röhling H, Mansow-Model S, Müller-Werdan U. Using new camera-based technologies for gait analysis in older adults in comparison to the established GAITRite system. *Sensors (Basel)* 2019 Dec;20(1):125 [FREE Full text] [doi: [10.3390/s20010125](https://doi.org/10.3390/s20010125)] [Medline: [31878177](https://pubmed.ncbi.nlm.nih.gov/31878177/)]
17. Mancini M, Salarian A, Carlson-Kuhta P, Zampieri C, King L, Chiari L, et al. ISway: a sensitive, valid and reliable measure of postural control. *J Neuroeng Rehabil* 2012 Aug;9:59 [FREE Full text] [doi: [10.1186/1743-0003-9-59](https://doi.org/10.1186/1743-0003-9-59)] [Medline: [22913719](https://pubmed.ncbi.nlm.nih.gov/22913719/)]
18. Kressig RW, Beauchet O, European GAITRite® Network Group. Guidelines for clinical applications of spatio-temporal gait analysis in older adults. *Aging Clin Exp Res* 2006 Apr;18:174-176. [doi: [10.1007/BF03327437](https://doi.org/10.1007/BF03327437)] [Medline: [16702791](https://pubmed.ncbi.nlm.nih.gov/16702791/)]
19. Hinton DC, Cheng Y, Paquette C. Everyday multitasking habits: university students seamlessly text and walk on a split-belt treadmill. *Gait Posture* 2018 Jan;59:168-173. [doi: [10.1016/j.gaitpost.2017.10.011](https://doi.org/10.1016/j.gaitpost.2017.10.011)] [Medline: [29032000](https://pubmed.ncbi.nlm.nih.gov/29032000/)]
20. Guffey K, Regier M, Mancinelli C, Pergami P. Gait parameters associated with balance in healthy 2- to 4-year-old children. *Gait Posture* 2016 Jan;43:165-169 [FREE Full text] [doi: [10.1016/j.gaitpost.2015.09.017](https://doi.org/10.1016/j.gaitpost.2015.09.017)] [Medline: [26439183](https://pubmed.ncbi.nlm.nih.gov/26439183/)]
21. Sunderland KM, Beaton D, Fraser J, Kwan D, McLaughlin PM, Montero-Odasso M, ONDRI Investigators, et al. The utility of multivariate outlier detection techniques for data quality evaluation in large studies: an application within the ONDRI project. *BMC Med Res Methodol* 2019 May;19:102 [FREE Full text] [doi: [10.1186/s12874-019-0737-5](https://doi.org/10.1186/s12874-019-0737-5)] [Medline: [31092212](https://pubmed.ncbi.nlm.nih.gov/31092212/)]
22. Martindale CF, Roth N, Gasner H, List J, Regensburger M, Eskofier BM, et al. Technical validation of an automated mobile gait analysis system for hereditary spastic paraplegia patients. *IEEE J Biomed Health Inform* 2020 May;24(5):1490-1499. [doi: [10.1109/JBHI.2019.2937574](https://doi.org/10.1109/JBHI.2019.2937574)] [Medline: [31449035](https://pubmed.ncbi.nlm.nih.gov/31449035/)]

23. Rampp A, Barth J, Schüle S, Gaßmann KG, Klucken J, Eskofier BM. Inertial sensor-based stride parameter calculation from gait sequences in geriatric patients. *IEEE Trans Biomed Eng* 2015 Apr;62(4):1089-1097. [doi: [10.1109/TBME.2014.2368211](https://doi.org/10.1109/TBME.2014.2368211)] [Medline: [25389237](https://pubmed.ncbi.nlm.nih.gov/25389237/)]
24. Grobelny A, Behrens JR, Mertens S, Otte K, Mansow-Model S, Krüger T, et al. Maximum walking speed in multiple sclerosis assessed with visual perceptive computing. *PLoS One* 2017 Dec;12(12):e0189281 [FREE Full text] [doi: [10.1371/journal.pone.0189281](https://doi.org/10.1371/journal.pone.0189281)] [Medline: [29244874](https://pubmed.ncbi.nlm.nih.gov/29244874/)]
25. Kroneberg D, Elshehabi M, Meyer A, Otte K, Doss S, Paul F, et al. Less is more—estimation of the number of strides required to assess gait variability in spatially confined settings. *Front Aging Neurosci* 2018 Jan;10:435 [FREE Full text] [doi: [10.3389/fnagi.2018.00435](https://doi.org/10.3389/fnagi.2018.00435)] [Medline: [30719002](https://pubmed.ncbi.nlm.nih.gov/30719002/)]
26. Behrens J, Pfüller C, Mansow-Model S, Otte K, Paul F, Brandt AU. Using perceptive computing in multiple sclerosis—the Short Maximum Speed Walk test. *J Neuroeng Rehabil* 2014 May;11:89 [FREE Full text] [doi: [10.1186/1743-0003-11-89](https://doi.org/10.1186/1743-0003-11-89)] [Medline: [24886525](https://pubmed.ncbi.nlm.nih.gov/24886525/)]
27. Behrens JR, Mertens S, Krüger T, Grobelny A, Otte K, Mansow-Model S, et al. Validity of visual perceptive computing for static posturography in patients with multiple sclerosis. *Mult Scler* 2016 Oct;22(12):1596-1606. [doi: [10.1177/1352458515625807](https://doi.org/10.1177/1352458515625807)] [Medline: [26814201](https://pubmed.ncbi.nlm.nih.gov/26814201/)]
28. Drebing D, Rasche L, Kroneberg D, Althoff P, Bellmann-Strobl J, Weygandt M, et al. Association between fatigue and motor exertion in patients with multiple sclerosis—a prospective study. *Front Neurol* 2020 Apr;11:208 [FREE Full text] [doi: [10.3389/fneur.2020.00208](https://doi.org/10.3389/fneur.2020.00208)] [Medline: [32351439](https://pubmed.ncbi.nlm.nih.gov/32351439/)]
29. Veauthier C, Ryczewski J, Mansow-Model S, Otte K, Kayser B, Glos M, et al. Contactless recording of sleep apnea and periodic leg movements by nocturnal 3-D-video and subsequent visual perceptive computing. *Sci Rep* 2019 Nov;9:16812 [FREE Full text] [doi: [10.1038/s41598-019-53050-3](https://doi.org/10.1038/s41598-019-53050-3)] [Medline: [31727918](https://pubmed.ncbi.nlm.nih.gov/31727918/)]
30. Otte K, Ellermeyer T, Suzuki M, Röhling HM, Kuroiwa R, Cooper G, et al. Cultural bias in motor function patterns: potential relevance for predictive, preventive, and personalized medicine. *EPMA J* 2021 Mar;12(1):91-101 [FREE Full text] [doi: [10.1007/s13167-021-00236-3](https://doi.org/10.1007/s13167-021-00236-3)] [Medline: [33782636](https://pubmed.ncbi.nlm.nih.gov/33782636/)]
31. Kurtzke JF. Rating neurologic impairment in multiple sclerosis: an expanded disability status scale (EDSS). *Neurology* 1983 Nov;33(11):1444-1452. [doi: [10.1212/WNL.33.11.1444](https://doi.org/10.1212/WNL.33.11.1444)] [Medline: [6685237](https://pubmed.ncbi.nlm.nih.gov/6685237/)]
32. Lachat E, Macher H, Landes T, Grussenmeyer P. Assessment and calibration of a RGB-D Camera (Kinect v2 Sensor) towards a potential use for close-range 3D modeling. *Remote Sens* 2015 Oct;7(10):13070-13097. [doi: [10.3390/rs71013070](https://doi.org/10.3390/rs71013070)]
33. Ramsperger R, Meckler S, Heger T, van Uem J, Hucker S, Braatz U, SENSE-PARK study team. Continuous leg dyskinesia assessment in Parkinson's disease—clinical validity and ecological effect. *Parkinsonism Relat Disord* 2016 May;26:41-46. [doi: [10.1016/j.parkreldis.2016.02.007](https://doi.org/10.1016/j.parkreldis.2016.02.007)] [Medline: [26952699](https://pubmed.ncbi.nlm.nih.gov/26952699/)]
34. Larsson J, Ekvall Hansson E, Miller M. Increased double support variability in elderly female fallers with vestibular asymmetry. *Gait Posture* 2015 Mar;41(3):820-824. [doi: [10.1016/j.gaitpost.2015.02.019](https://doi.org/10.1016/j.gaitpost.2015.02.019)] [Medline: [25800649](https://pubmed.ncbi.nlm.nih.gov/25800649/)]
35. Dusing SC, Thorpe DE. A normative sample of temporal and spatial gait parameters in children using the GAITRite® electronic walkway. *Gait Posture* 2007 Jan;25(1):135-139. [doi: [10.1016/j.gaitpost.2006.06.003](https://doi.org/10.1016/j.gaitpost.2006.06.003)] [Medline: [16875823](https://pubmed.ncbi.nlm.nih.gov/16875823/)]
36. Schmitz-Hübsch T, Brandt AU, Pfueller C, Zange L, Seidel A, Kühn AA, et al. Accuracy and repeatability of two methods of gait analysis – GaitRite™ und Mobility Lab™ – in subjects with cerebellar ataxia. *Gait Posture* 2016 Jul;48:194-201. [doi: [10.1016/j.gaitpost.2016.05.014](https://doi.org/10.1016/j.gaitpost.2016.05.014)] [Medline: [27289221](https://pubmed.ncbi.nlm.nih.gov/27289221/)]
37. Wong JS, Jasani H, Poon V, Inness EL, McIlroy WE, Mansfield A. Inter- and intra-rater reliability of the GAITRite system among individuals with sub-acute stroke. *Gait Posture* 2014 May;40(1):259-261. [doi: [10.1016/j.gaitpost.2014.02.007](https://doi.org/10.1016/j.gaitpost.2014.02.007)] [Medline: [24630463](https://pubmed.ncbi.nlm.nih.gov/24630463/)]
38. Tewarie P, Balk L, Costello F, Green A, Martin R, Schippling S, et al. The OSCAR-IB consensus criteria for retinal OCT quality assessment. *PLoS One* 2012 Apr;7(4):e34823 [FREE Full text] [doi: [10.1371/journal.pone.0034823](https://doi.org/10.1371/journal.pone.0034823)] [Medline: [22536333](https://pubmed.ncbi.nlm.nih.gov/22536333/)]
39. Cruz-Herranz A, Balk LJ, Oberwahrenbrock T, Saidha S, Martinez-Lapiscina EH, Lagreze WA, IMSVISUAL consortium. The APOSTEL recommendations for reporting quantitative optical coherence tomography studies. *Neurology* 2016 Jun;86(24):2303-2309 [FREE Full text] [doi: [10.1212/WNL.0000000000002774](https://doi.org/10.1212/WNL.0000000000002774)] [Medline: [27225223](https://pubmed.ncbi.nlm.nih.gov/27225223/)]
40. Traboulsee A, Simon JH, Stone L, Fisher E, Jones DE, Malhotra A, et al. Revised recommendations of the consortium of MS Centers Task Force for a standardized MRI protocol and clinical guidelines for the diagnosis and follow-up of multiple sclerosis. *AJNR Am J Neuroradiol* 2016 Mar;37(3):394-401 [FREE Full text] [doi: [10.3174/ajnr.A4539](https://doi.org/10.3174/ajnr.A4539)] [Medline: [26564433](https://pubmed.ncbi.nlm.nih.gov/26564433/)]
41. Motamedi S, Gawlik K, Ayadi N, Zimmermann HG, Asseger S, Bereuter C, et al. Normative data and minimally detectable change for inner retinal layer thicknesses using a semi-automated OCT image segmentation pipeline. *Front Neurol* 2019 Nov;10:1117 [FREE Full text] [doi: [10.3389/fneur.2019.01117](https://doi.org/10.3389/fneur.2019.01117)] [Medline: [31824393](https://pubmed.ncbi.nlm.nih.gov/31824393/)]
42. Chien C, Juenger V, Scheel M, Brandt AU, Paul F. Considerations for mean upper cervical cord area implementation in a longitudinal MRI setting: methods, interrater reliability, and MRI quality control. *AJNR Am J Neuroradiol* 2020 Feb;41(2):343-350. [doi: [10.3174/ajnr.A6394](https://doi.org/10.3174/ajnr.A6394)] [Medline: [31974079](https://pubmed.ncbi.nlm.nih.gov/31974079/)]

Abbreviations

GUI: graphical user interface

IMA: instrumented motion analysis

MS: multiple sclerosis

PASS-MS: Perceptive Assessment in Multiple Sclerosis (ie, name of short motor assessment battery recorded with Motognosis Labs)

POCO: Postural Control

POCO-DUAL: Postural Control with Dual Task

QC: quality control

SAS: Stand Up and Sit Down

SCSW: Short Comfortable Speed Walk

SIP: Stepping in Place

SLW: Short Line Walk

SMSW: Short Maximum Speed Walk

SOP: standard operating procedure

Edited by A Kushniruk; submitted 29.12.20; peer-reviewed by C Mazza, E Ravera; comments to author 16.03.21; revised version received 02.05.21; accepted 07.12.21; published 01.04.22.

Please cite as:

Röhling HM, Althoff P, Arsenova R, Drebinger D, Gigengack N, Chorschew A, Kroneberg D, Rönnefarth M, Ellermeyer T, Rosenkranz SC, Heesen C, Behnia B, Hirano S, Kuwabara S, Paul F, Brandt AU, Schmitz-Hübsch T

Proposal for Post Hoc Quality Control in Instrumented Motion Analysis Using Markerless Motion Capture: Development and Usability Study

JMIR Hum Factors 2022;9(2):e26825

URL: <https://humanfactors.jmir.org/2022/2/e26825>

doi: [10.2196/26825](https://doi.org/10.2196/26825)

PMID: [35363150](https://pubmed.ncbi.nlm.nih.gov/35363150/)

©Hanna Marie Röhling, Patrik Althoff, Radina Arsenova, Daniel Drebinger, Norman Gigengack, Anna Chorschew, Daniel Kroneberg, Maria Rönnefarth, Tobias Ellermeyer, Sina Cathérine Rosenkranz, Christoph Heesen, Behnoush Behnia, Shigeki Hirano, Satoshi Kuwabara, Friedemann Paul, Alexander Ulrich Brandt, Tanja Schmitz-Hübsch. Originally published in JMIR Human Factors (<https://humanfactors.jmir.org>), 01.04.2022. This is an open-access article distributed under the terms of the Creative Commons Attribution License (<https://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Human Factors, is properly cited. The complete bibliographic information, a link to the original publication on <https://humanfactors.jmir.org>, as well as this copyright and license information must be included.

Original Paper

Evaluation of the First Year(s) of Physicians Collaboration on an Interdisciplinary Electronic Consultation Platform in the Netherlands: Mixed Methods Observational Study

Sanne M Sanavro¹, MD; Henk van der Worp^{1*}, PhD; Danielle Jansen^{1*}, PhD; Paul Koning^{2*}, MD; Marco H Blanker^{1*}, MD, PhD; Prisma Platform Study Group^{3*}

¹Department of General Practice and Elderly Care Medicine, University of Groningen, Groningen, Netherlands

²Prisma, Siilo Holding BV, Amsterdam, Netherlands

³See Acknowledgments

*these authors contributed equally

Corresponding Author:

Sanne M Sanavro, MD

Department of General Practice and Elderly Care Medicine

University of Groningen

PO Box 196

Groningen, 9700 AD

Netherlands

Phone: 31 50 3616731

Fax: 31 50 3632964

Email: sannesanavro@hotmail.com

Abstract

Background: Complexity of health problems and aging of the population create an ongoing burden on the health care system with the general practitioner (GP) being the gatekeeper in primary care. In GPs daily practice, collaboration with specialists and exchange of knowledge from the secondary care play a crucial role in this system. Communication between primary and secondary care has shortcomings for health care workers that want to practice sustainable patient-centered health care. Therefore, a new digital interactive platform was developed: Prisma.

Objective: This study aims to describe the development of a digital consultation platform (Prisma) to connect GPs with hospital specialists via the Siilo application and to evaluate the first year of use, including consultations, topic diversity, and number of participating physicians.

Methods: We conducted a mixed methods observational study, analyzing qualitative and quantitative data for cases posted on the platform between June 2018 and May 2020. Any GP can post questions to an interdisciplinary group of secondary care specialists, with the platform designed to facilitate discussion and knowledge exchange for all users.

Results: In total, 3674 cases were posted by 424 GPs across 16 specialisms. Most questions and answers concerned diagnosis, nonmedical treatment, and medication. Mean response time was 76 minutes (range 44-252). An average of 3 users engaged with each case (up to 7 specialists). Almost half of the internal medicine cases received responses from at least two specialisms in secondary care, contrasting with about one-fifth for dermatology. Of note, the growth in consultations was steepest for dermatology.

Conclusions: Digital consultations offer the possibility for GPs to receive quick responses when seeking advice. The interdisciplinary approach of Prisma creates opportunities for digital patient-centered networking.

(*JMIR Hum Factors* 2022;9(2):e33630) doi:[10.2196/33630](https://doi.org/10.2196/33630)

KEYWORDS

primary care; digital consultation; interdisciplinary; specialist care

Introduction

In the Dutch health care system, general practitioners (GPs) have a coordinating role as generalists, functioning as gatekeepers to secondary care. This model requires that patients initially consult a GP who provides expert generalist medical care for their health care problem and considers the need for referral to more specialist care.

Unfortunately, pressures on the health care system have increased due to the growth in both the chronicity and the complexity of health problems [1,2]. Although GPs care for over 95% of medical problems that present during consultations, referral to secondary care has also increased, resulting in greater health care costs and growing waiting lists [3,4]. These issues can be addressed by providing GPs with closer support from secondary care, assuming there are effective routes for knowledge exchange [5-9]. However, the most commonly used tools for communication between primary and secondary care have important shortcomings. For example, GPs and hospital specialists are often mutually unavailable at the same time, meaning that telephone conversations can be interruptive. Whereas e-consultations may solve the problem of asynchronous availability, they are limited by being monodisciplinary, one-on-one, and mostly noninteractive [6,10-13]. Digital response times may also vary by specialism. By contrast, team-based case collaboration on a patient-centered network of health care professionals could facilitate communication and knowledge transfer [14-16]. The secure Siilo app offers a useful platform to host such a service [17,18].

In this study, we describe the development of the Dutch Prisma platform within the secure Siilo app and evaluate the usage and consultations in the first 2 years since its introduction, including the diversity of topics and number of physicians involved.

Methods

Study Design

We performed a retrospective mixed methods study using quantitative information from the Prisma platform and a qualitative evaluation of consecutive cases posted on the platform from its inception in July 2018 to May 2020.

The Prisma Platform

The Prisma platform initially facilitated digital interprofessional consultation for patients with orthopedic problems, but more recently, it has expanded to include other specialties. GPs with full access to the closed digital environment of the platform generate cases by providing anonymized patient information with a question. All GPs and specialist users are connected in so-called tiles by specialty (eg, orthopedics, internal medicine, palliative care) to facilitate engagement by consultants with complementary expertise (eg, rheumatologists, orthopedic surgeons, sports medicine physicians, and radiologists participate via the orthopedics tile). All users can engage with each tile and upload attachments or links to relevant information, such as laboratory results, pictures, or guidelines. The main language used on the platform is Dutch.

Two GP groups are active on this platform: 1 with full access (able to generate cases and respond to others) and 1 with a read-only account. Specialists participated voluntarily; separate from their hospital work and without reimbursement for their activity on the platform. Because they were not reimbursed, the number of GPs was limited during the development phase to avoid overloading the specialists. All users, both GPs and specialists, were located in various regions of the Netherlands. Specialists preferably respond within 24-48 hours by answering questions, seeking more information, or engaging in discussion. All GPs with access to the platform can read and respond to posted cases. In this way, the platform allows for a dynamic exchange of information and learning to support the GP in daily practice. Throughout the process, the GP remains responsible for the care provided to the patient and will decide, in consultation with the patient, how to proceed with further treatment.

Data Collection

A data analyst at Siilo provided pseudonymized details for all consecutive cases, replacing usernames with a job title and a number (eg, GP-1, GP-2, neurologist-1). Each post was summarized as a user code, timestamp, and verbatim transcript, and these were grouped by case for each tile. Data were analyzed qualitatively and quantitatively. As we performed a retrospective descriptive study, we did not predefine our sample size.

Qualitative Analysis

Text files were imported into the Atlas.ti program [19] for qualitative assessment by a research group comprising 20 senior medical students (coders) supervised by an internist (SS), a medical sociologist (DJ), a GP epidemiologist (MB), and a senior researcher (HW). The Prisma affiliate (PK) was not involved in this phase.

We used a predefined coding tree to structure the qualitative assessment (Multimedia Appendix 1). Before applying this to all cases, a random sample of 10 cases was initially coded by all coders. The results of this preliminary coding were then checked in pairs and discussed in 5 subgroups with 2 supervisors. Coders were actively invited to discuss the applicability of codes and to add new codes if needed. After this, coders were grouped by tile and at least 50 cases per tile were coded in duplicate with mutual blinding. This was followed by group discussion in consensus meetings per subgroup, after which the remaining cases were coded.

The coding tree comprised the following: basic patient characteristics, such as age, gender, and comorbidity; the topic of the question; and both the type of question and the type of answer (eg, diagnostic, therapeutic, or referral for both). Codes for symptoms and diseases followed the International Classification in Primary Care (ICPC) [20], with multiple codes permitted.

Quantitative Analysis

All codes were imported into IBM SPSS (IBM Corp.) for quantitative analysis. We merged the 16 tiles into 5 categories based on similarities and group sizes: "internal medicine" included internal medicine, infectious disease, palliative care,

and medically unexplained physical symptoms; “observation” included gastroenterology, neurology, pulmonology, rheumatology, and cardiology; “surgical” included orthopedics, urology, traumatology, and ear, nose, and throat disease; “female/child” included gynecology and pediatrics; and “dermatology” as a single category. The tile for psychiatry was analyzed and published separately and is therefore excluded from this analysis [21].

An overview of activity on the platform is displayed by plotting the number of GPs (active users and read-only accounts) and the number of cases against time. We estimated the number of users, number of specialisms, number of specialists, and the response time for each case based on user codes and timestamps, and we analyzed the code frequencies for age, gender, case topic (based on the ICPC code), question type, and answer type for each category. Descriptive data were presented as percentages of all cases or as means and SDs. Finally, we used a Sankey diagram to show the linkage between questions and answers.

Results

Descriptive Data

The data set started with 25,954 messages for 4013 cases; of these, 1872 messages were excluded for 339 cases. First, we excluded 292 cases because of data extraction errors ($n=34$),

small size, and difficulty to categorize within groups (geriatrics, $n=5$; ophthalmology, $n=40$) and because they were already analyzed in a separate study (psychiatry, $n=213$) [21]. Next, we divided the data within the research team and analyzed the 3721 cases. We excluded another 47 cases because of wrong tile placement ($n=19$), double case placement ($n=10$), technical errors ($n=8$), not coded ($n=7$), withdrawal by GP ($n=2$), missing ($n=1$) (Multimedia Appendix 2). The 3674 included cases were posted by 424 different GPs (median 9 cases per GP), for whom 97 (22.9%) first posts were in response to another case and 327 (77.1%) posts were for new cases.

Growth of the Prisma platform over time is shown as the number of GPs (active users and read-only accounts; Figure 1), the total number of cases, and the number of cases per tile (Figure 2). The number of cases per category was 677 for internal, 674 for observation, 860 for surgical, 875 for female/child, and 588 for dermatology. Figures 3 and 4 show the number of specialists and specialisms involved per tile category, respectively. For all categories, except dermatology (196/588, 33.3%), most cases included more than 2 users per case. For the internal, observation, and surgical categories, 3 or more specialisms were involved per case in 46.6% (317/680), 32.3% (217/672), and 40.7% (350/860), respectively. In the internal and observation categories, 4 or more health care professionals were engaged per case in 57.2% (389/680) and 54.0% (363/672), respectively.

Figure 1. Platform use; number of active and read-only GPs on the platform. GP: general practitioner.

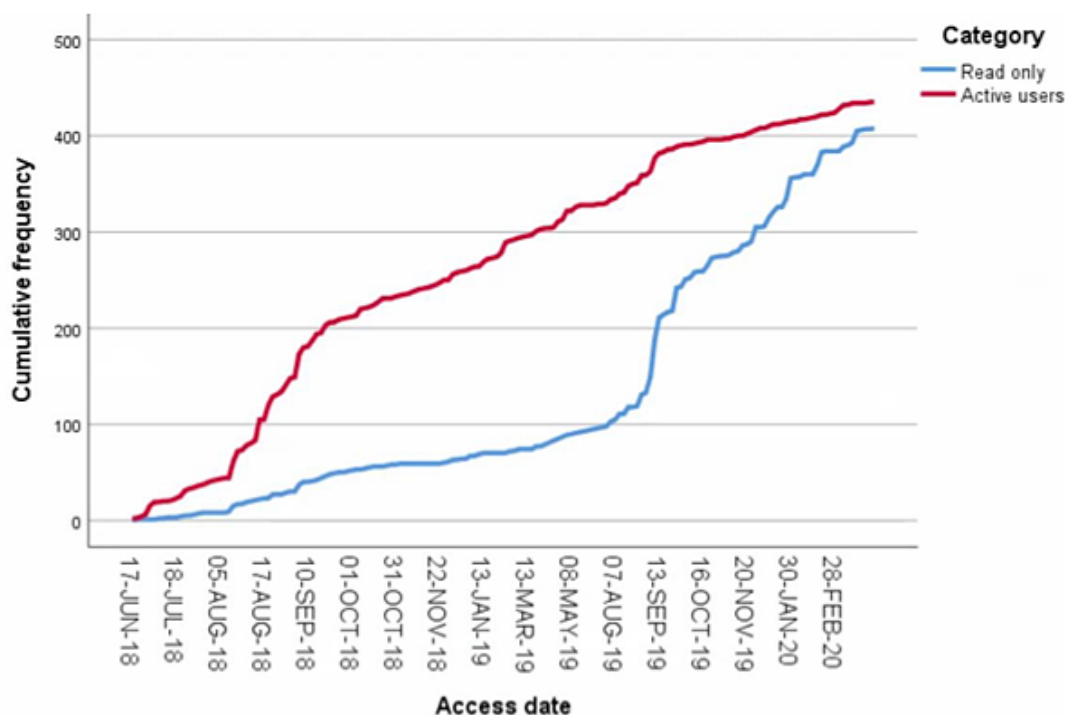


Figure 2. Overall cases of network activity and network activity by tile category. ENT: ear, nose, throat; GYN: gynaecology; MUPS: medically unexplained physical symptoms; PAL: palliative care; UROL: urology.

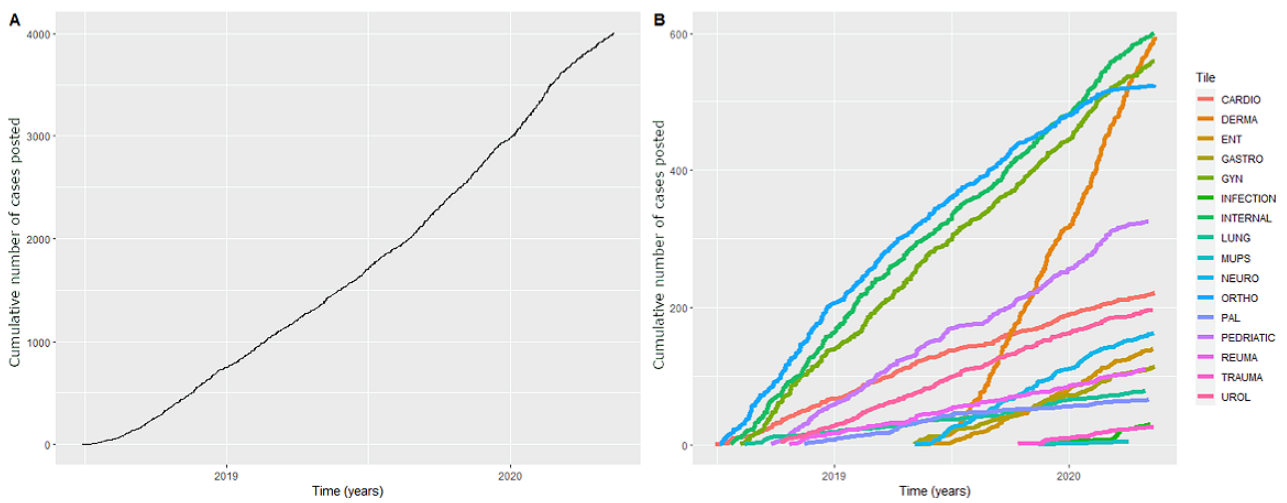


Figure 3. Number of users involved per case. Data are illustrated in 5 tile categories.

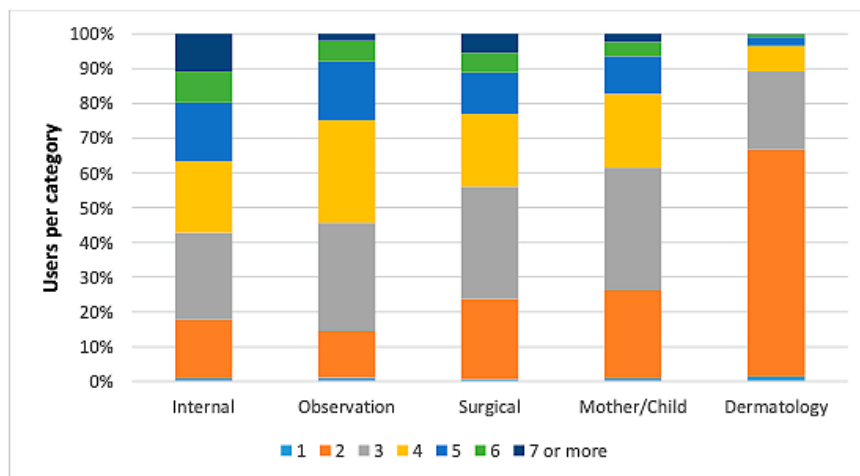
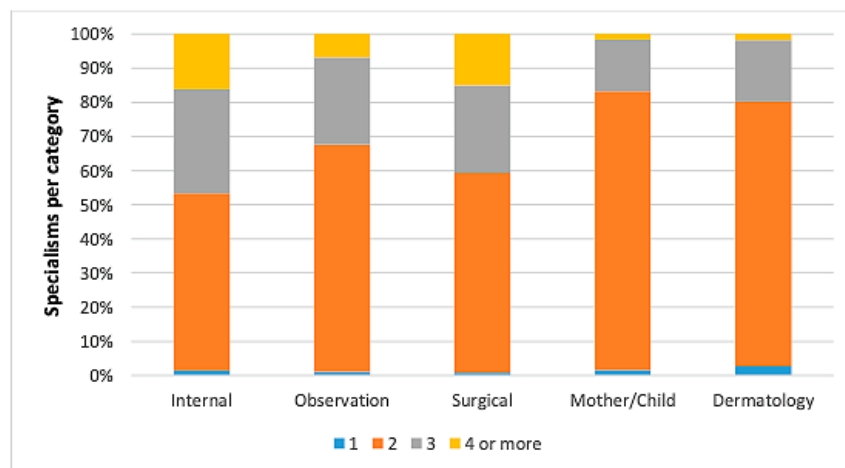


Figure 4. Number of specialisms involved per case. Data are illustrated in 5 tile categories.



Case characteristics are presented in [Multimedia Appendix 3](#). No answer was given for 35 cases, with the median time to first response being 76 minutes (IQR 17-320) for the other cases. The shortest response time was seen in the surgery category (median 44 minutes) and the longest was in the dermatology tile (median 252 minutes). Overall, 3508/3674 (95.48%) cases contained specific patient information or patient-specific questions, with the remaining 166 (4.52%) cases including questions that were not specific to the patient. Slightly more than half of all queries concerned females (1948/3674, 53.02%), except for those in the surgical tile where there was a slight male majority (437/860, 50.8%). GPs did not report gender in 8.92% (313/3508) of the patient-specific cases. They also posted a question about more than 1 patient in 4 cases (eg, family members or several patients with the same complaint). Patient age ranged from newborn to 101 years (mean 39.9 years) and the mean age differed by tile category. The GP did not report age for 701 cases.

Topics discussed covered the full range of ICPC codes ([Multimedia Appendix 4](#)). The 3 main topics by ICPC code were in the skin, musculoskeletal, and general symptom domains.

Type of Questions and Answers

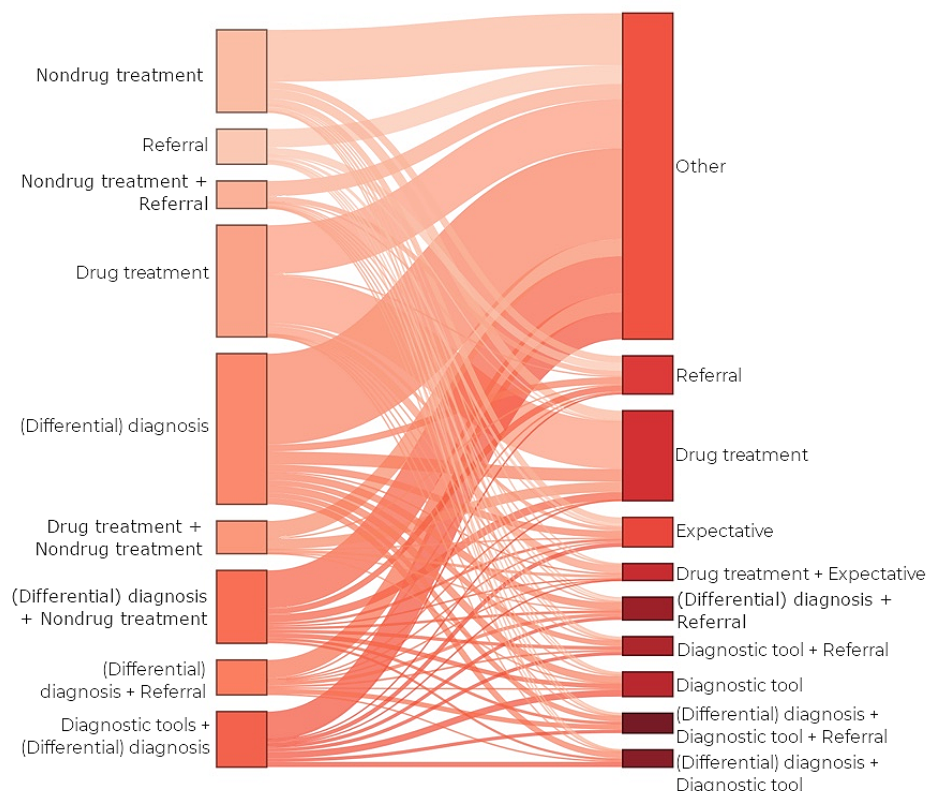
Among the 3674 cases, we identified 6691 different questions (mean 1.8 per case) and 10,922 answers (mean 3.03 per case).

[Multimedia Appendix 5](#) shows the type of question and answers posted.

Questions concerned (differential) diagnosis in 50.90% (1870/3674), appropriate nondrug treatment in 33.15% (1218/3674), and drug treatment in 27.60% (1014/3674). It was notable that the focus of questions differed between tile categories. Most concerned diagnosis in the internal (358/677, 52.9%), observatory (361/674, 53.6%), and dermatology (424/588, 72.1%) categories; most concerned treatment in the surgical category (431/860, 50.1%); and most concerned medication in the female/child category (378/875, 43.2%).

The Sankey diagram in [Figure 5](#) illustrates the dynamics between the type of question and the type of answer. We have illustrated only the 9 most common combinations (used more than 100 times), including any other answer type or combination in the “other” group. Consistent with the type of question asked by GPs, most answers concerned (differential) diagnosis, which was often combined with responses about referral, further diagnostics, or a combination of these 3 responses. However, the type of question posed by GPs did not always lead to answers within the same topic, such as questions about referral often leading to advice about how to proceed (eg, perform further diagnostics and refer, GP-based follow-up, or start therapy and refer). In this way, one can see that simple referral questions can lead to varied advice possibilities ([Multimedia Appendices 6-8](#)).

Figure 5. Sankey diagram of dynamics from questions to answers, 9 largest groups.



Discussion

Principal Findings

This mixed methods study has shown the growth and evolution of a digital interdisciplinary consultation platform over almost 2 years. Posted questions not only covered a broad spectrum of the population by age and sex but also covered a wide variety of specialist topics. Of note, there was a steep increase in the number of cases for dermatology, which could be explained by existing familiarity with tele-dermatology in Dutch primary care [10] or potentially highlight a practice weakness among GPs.

In most cases, 2 or more users engaged with the GP who initiated the question. An exception to this was the dermatology tile, in which it was typical for only 1 other user to respond. The number of involved specialisms also differed between tiles, being largest for internal medicine. This illustrates a novelty of this approach compared with other consultation formats where a GP only has contact with 1 medical specialist. This approach is in line with the future vision to build primary and secondary care networks around the patient [16,22,23].

The short response times suggest that the Prisma platform facilitates rapid and efficient consultation. This contrasts with telephone consultations, which are often hampered by mutual unavailability. Our data indicate that answers are given to most questions by the end of a GP's working day so that patient care is not delayed for more than a few hours.

Although it is difficult to compare our study with previous studies because of the difference in design of the platform that was analyzed, the time response outcomes are superior to those in previous studies [4,6,11,24]. It should be considered that they may reflect a precursor effect of enthusiasm among engaged specialists.

The differences in question type between tile categories may indicate differences in work content. Internal medicine, observation, and dermatology focused on diagnosis; surgery focused on treatment; and female/child focused on medication. An alternative hypothesis could be that different specialisms have specific needs of GPs in the treatment process.

The Sankey graph in [Figure 5](#) and [Multimedia Appendices 6-8](#) illustrates the dynamics between questions asked by GPs and answers given by specialists. The large number of questions related to diagnosis had multiple combinations with other questions, reflecting the complexity of evaluation (eg, when the diagnosis is unclear, the next step is also uncertain). Overall, (differential) diagnosis was the most frequently used theme, but this does not appear as a separate group in the graph because it was mostly used in combination with other themes. In comparison to this, questions on medication had most single questions and a clear dynamic to single answers.

The dynamics on referral questions are also interesting, with only a minority of questions receiving a single answer about referral. For example, we found combinations of advice for additional diagnostics in primary care or advice to refer with explanations about diagnosis. We hypothesize that medical

specialists used this platform not only to ensure adequate referral but also to share knowledge. There was also a difference between referral questions and answers: not all questions about referral led to answers about referral, and vice versa (ie, referral advice was sometimes given without a specific request).

We found similarities and differences when comparing our findings with the limited amount of preceding research on electronic consultations [15]. In this earlier research, most questions for hematology and rheumatology concerned diagnosis, while questions in the infectious disease and dermatology categories typically concerned therapy. Another research focusing specifically on internal medicine in a hospital in Netherlands involved one-on-one electronic consultations, and revealed "diagnostic tools" to be the most common answer [6].

Limitations

First, the large sample size and categorization means that a more detailed analysis by specialty is missing in this study. Second, because structure was lacking in the questions posted by GPs, complete data on patient characteristics cannot be guaranteed; however, this did not impair the content analysis. Third, text coding was done by 20 different coders, which might have resulted in interobserver variations in interpretation, despite our efforts to minimize this as much as possible through teamwork. Finally, the data in this analysis were observational in nature, preventing us from making firm conclusions on either observed correlations or patient outcomes.

Future Research

This evaluation focused on the activities of health care professionals, but to date, the patient perspective has not been analyzed. Although the platform performs well in supporting the needs of the GP for further assessment, treatment, and when needed, more appropriate referral to specialists, we do not know how these relate to needs, experiences, and outcomes in patient cohorts. To generate and implement a novel health care collaboration on a large scale, time and cost-efficiency calculations will also be indispensable [25]. In our study, the response time was more rapid than previously reported for e-consultations [6,24], which have already been shown to reduce not only waiting times for GPs and patients but also costs for patients and waiting lists for hospitals [26]. We are currently conducting a stepped-wedge randomized controlled trial to evaluate the impact of the Prisma platform on patient outcomes and referrals to specialists.

Concerning the content of questions posted on the Prisma platform, an in-depth analysis could still be interesting and useful. Gaps in support for GPs could be uncovered by exploring diagnostic uncertainties (between noncomplex symptoms that meet ICPC diagnostic criteria and practice guidelines), common reasons for referral, and the impact of regional agreements [27]. It is possible that these gaps could be filled by creating a database of information collected on the platform. This could facilitate GPs to ask questions and search for possible answers based on prior responses.

Conclusion

This observational research shows that a new digital platform facilitated rapid and interactive communication between GPs and specialists for nonurgent questions. This platform is clearly

distinguished from one-to-one consultations by facilitating the involvement of multiple physicians. The platform supports the transfer of knowledge from medical specialists to GPs while allowing different viewpoints from relevant experts.

Acknowledgments

We are grateful to the Department of General Practice, University Medical Centre Groningen, and to Prisma for providing technical support. Prisma Platform Study Group: Ineke Knijp, Ruben BR de Boer, Lisa Havinga, Iris R Vroom, Esmee L van der Geest, Maaïke A Hulshof, Daphne BM Visser, Frank J Dorgelo, Ivar Maatje, Hugo Quaedvlieg, Lianne Sijbring, Anne van der Meer, Esmée Fath el Bab, Lucinda E Haaze, Willeke Schelhaas, Susan JM Bergamin, Kimberly Boerma, Anne J Lammers, Dajana Erceg, Tamara Schouten, and Steven N Koning. We thank Dr Robert Sykes for providing technical editing services for the final drafts of this manuscript.

Conflicts of Interest

PK is the founder of Prisma and provided the text messages of the platform for this research, but was not involved in the data analysis process of this work.

Multimedia Appendix 1

Code tree (in Dutch).

[[PDF File \(Adobe PDF File\), 165 KB - humanfactors_v9i2e33630_app1.pdf](#)]

Multimedia Appendix 2

Flowchart inclusion cases.

[[PNG File , 24 KB - humanfactors_v9i2e33630_app2.png](#)]

Multimedia Appendix 3

Characteristics of cases posted in different tiles.

[[DOC File , 37 KB - humanfactors_v9i2e33630_app3.doc](#)]

Multimedia Appendix 4

International Classification in Primary Care codes used in cases.

[[DOC File , 47 KB - humanfactors_v9i2e33630_app4.doc](#)]

Multimedia Appendix 5

Type of questions and type of answers.

[[DOC File , 53 KB - humanfactors_v9i2e33630_app5.doc](#)]

Multimedia Appendix 6

Sankey diagram for differential diagnostic questions. DT: drug treatment; NDT: nondrugs treatment.

[[PNG File , 1867 KB - humanfactors_v9i2e33630_app6.png](#)]

Multimedia Appendix 7

Sankey diagram for drug treatment questions.

[[PNG File , 2742 KB - humanfactors_v9i2e33630_app7.png](#)]

Multimedia Appendix 8

Sankey diagram for referral questions.

[[PNG File , 390 KB - humanfactors_v9i2e33630_app8.png](#)]

References

1. van Oostrom SH, Picavet HSJ, de Bruin SR, Stirbu I, Korevaar JC, Schellevis FG, et al. Multimorbidity of chronic diseases and health care utilization in general practice. *BMC Fam Pract* 2014 Apr 07;15:61 [[FREE Full text](#)] [doi: [10.1186/1471-2296-15-61](https://doi.org/10.1186/1471-2296-15-61)] [Medline: [24708798](https://pubmed.ncbi.nlm.nih.gov/24708798/)]

2. Schäfer WLA, Boerma W, Spreeuwenberg P, Schellevis F, Groenewegen P. Two decades of change in European general practice service profiles: conditions associated with the developments in 28 countries between 1993 and 2012. *Scand J Prim Health Care* 2016;34(1):97-110 [FREE Full text] [doi: [10.3109/02813432.2015.1132887](https://doi.org/10.3109/02813432.2015.1132887)] [Medline: [26862927](https://pubmed.ncbi.nlm.nih.gov/26862927/)]
3. Kroneman M, Boerma W, van den Berg M, Groenewegen P, de Jong J, van Ginneken E. Netherlands: Health System Review. *Health Syst Transit* 2016 Mar;18(2):1-240 [FREE Full text] [Medline: [27467715](https://pubmed.ncbi.nlm.nih.gov/27467715/)]
4. Liddy C, Moroz I, Afkham A, Keely E. Sustainability of a Primary Care-Driven eConsult Service. *Ann Fam Med* 2018 Mar;16(2):120-126 [FREE Full text] [doi: [10.1370/afm.2177](https://doi.org/10.1370/afm.2177)] [Medline: [29531102](https://pubmed.ncbi.nlm.nih.gov/29531102/)]
5. Berendsen A, Benneker W, Meyboom-de Jong B, Klazinga N, Schuling J. Motives and preferences of general practitioners for new collaboration models with medical specialists: a qualitative study. *BMC Health Serv Res* 2007 Jan 05;7:4 [FREE Full text] [doi: [10.1186/1472-6963-7-4](https://doi.org/10.1186/1472-6963-7-4)] [Medline: [17207278](https://pubmed.ncbi.nlm.nih.gov/17207278/)]
6. Muris D, Krekels M, Spreeuwenberg A, Blom M, Bergmans P, Cals JWL. General practitioners' use of internal medicine e-consultations. *Ned Tijdschr Geneesk* 2020 Feb 10;164:24-31. [Medline: [32186815](https://pubmed.ncbi.nlm.nih.gov/32186815/)]
7. Wagner EH, Glasgow RE, Davis C, Bonomi AE, Provost L, McCulloch D, et al. Quality Improvement in Chronic Illness Care: A Collaborative Approach. *The Joint Commission Journal on Quality Improvement* 2001 Feb;27(2):63-80. [doi: [10.1016/s1070-3241\(01\)27007-2](https://doi.org/10.1016/s1070-3241(01)27007-2)]
8. de Bever S, Bont J, Scherpier N. Strengthening general practice by extending specialty training? *Br J Gen Pract* 2019 Apr 25;69(682):222-223. [doi: [10.3399/bjgp19x702305](https://doi.org/10.3399/bjgp19x702305)]
9. Greenwood-Lee J, Jewett L, Woodhouse L, Marshall D. A categorisation of problems and solutions to improve patient referrals from primary to specialty care. *BMC Health Serv Res* 2018 Dec 20;18(1):986 [FREE Full text] [doi: [10.1186/s12913-018-3745-y](https://doi.org/10.1186/s12913-018-3745-y)] [Medline: [30572898](https://pubmed.ncbi.nlm.nih.gov/30572898/)]
10. Tensen E, van der Heijden JP, Jaspers M, Witkamp L. Two Decades of Tele dermatology: Current Status and Integration in National Healthcare Systems. *Curr Dermatol Rep* 2016;5:96-104 [FREE Full text] [doi: [10.1007/s13671-016-0136-7](https://doi.org/10.1007/s13671-016-0136-7)] [Medline: [27182461](https://pubmed.ncbi.nlm.nih.gov/27182461/)]
11. Liddy C, Moroz I, Mihan A, Nawar N, Keely E. A Systematic Review of Asynchronous, Provider-to-Provider, Electronic Consultation Services to Improve Access to Specialty Care Available Worldwide. *Telemed J E Health* 2019 Mar;25(3):184-198 [FREE Full text] [doi: [10.1089/tmj.2018.0005](https://doi.org/10.1089/tmj.2018.0005)] [Medline: [29927711](https://pubmed.ncbi.nlm.nih.gov/29927711/)]
12. Soriano Marcolino M, Minelli Figueira R, Pereira Afonso Dos Santos J, Silva Cardoso C, Luiz Ribeiro A, Alkmim M. The Experience of a Sustainable Large Scale Brazilian Telehealth Network. *Telemed J E Health* 2016 Nov;22(11):899-908 [FREE Full text] [doi: [10.1089/tmj.2015.0234](https://doi.org/10.1089/tmj.2015.0234)] [Medline: [27167901](https://pubmed.ncbi.nlm.nih.gov/27167901/)]
13. Kwok J, Olayiwola J, Knox M, Murphy E, Tuot D. Electronic consultation system demonstrates educational benefit for primary care providers. *J Telemed Telecare* 2017 Jun 14;24(7):465-472 [FREE Full text] [doi: [10.1177/1357633x17711822](https://doi.org/10.1177/1357633x17711822)]
14. McIntyre T, Kelly E, Clarke T, Green C. Design and implementation of an acute Trauma and Orthopaedic ePlatform (TOP) referral system utilising existing secure technology during the COVID-19 pandemic. *Bone & Joint Open* 2020 Jun 01;1(6):293-301 [FREE Full text] [doi: [10.1302/2633-1462.16.bjo-2020-0041.r1](https://doi.org/10.1302/2633-1462.16.bjo-2020-0041.r1)]
15. Waugh M, Voyles D, Thomas M. Telepsychiatry: Benefits and costs in a changing health-care environment. *Int Rev Psychiatry* 2015;27(6):558-568 [FREE Full text] [doi: [10.3109/09540261.2015.1091291](https://doi.org/10.3109/09540261.2015.1091291)] [Medline: [26586193](https://pubmed.ncbi.nlm.nih.gov/26586193/)]
16. Miller R, Scherpier N, van Amsterdam L, Guedes V, Pype P. Inter-professional education and primary care: EFPC position paper. *Prim Health Care Res Dev* 2019 Oct 04;20:E138 [FREE Full text] [doi: [10.1017/s1463423619000653](https://doi.org/10.1017/s1463423619000653)]
17. Security Whitepaper. Siilo. URL: <https://www.siilo.com/assets/downloads/Siilo-Security-Whitepaper.pdf> [accessed 2021-09-16]
18. Ezra O, Toren A, Tadmor O, Katorza E. Secure Instant Messaging Application in Prenatal Care. *J Med Syst* 2020 Feb 22;44(4):73 [FREE Full text] [doi: [10.1007/s10916-020-01552-3](https://doi.org/10.1007/s10916-020-01552-3)] [Medline: [32086609](https://pubmed.ncbi.nlm.nih.gov/32086609/)]
19. Friese S. Atlas.ti 8, Windows User Manual. Atlasti. 2015. URL: https://downloads.atlasti.com/docs/branding/atlasti_brochure_v9_EN_interactive_202110.pdf?_ga=2.83602454.1041389821.1647378949-1120635778.1647378949 [accessed 2022-03-15]
20. International Classification of Primary Care, Second Edition (ICPC-2). ICPC. 2014. URL: <https://www.transhis.nl/wp-content/uploads/2014/12/icpc-2-2pager-nederlands.pdf> [accessed 2022-03-15]
21. Bock NW, Wouters H, Lammers AJ, Blanker MH. Online Consultations Between General Practitioners and Psychiatrists in the Netherlands: A Qualitative Study. *Front Psychiatry* 2021;12:775738 [FREE Full text] [doi: [10.3389/fpsy.2021.775738](https://doi.org/10.3389/fpsy.2021.775738)] [Medline: [34803781](https://pubmed.ncbi.nlm.nih.gov/34803781/)]
22. Reddy S. Integrated health care: it's time for it to blossom. *Aust. Health Review* 2016;40(4):428-430 [FREE Full text] [doi: [10.1071/ah15135](https://doi.org/10.1071/ah15135)]
23. World Health Organization (WHO). WHO global strategy on people-centred and integrated health services. WHO. Geneva, Switzerland: WHO; 2015. URL: https://apps.who.int/iris/bitstream/handle/10665/155002/WHO_HIS_SDS_2015.6_eng.pdf?sequence=1&isAllowed=y [accessed 2022-03-20]
24. Ahmed S, Kelly Y, Behera T, Zelen M, Kuye I, Blakey R, et al. Utility, Appropriateness, and Content of Electronic Consultations Across Medical Subspecialties. *Annals of Internal Medicine* 2020 May 19;172(10):641-647 [FREE Full text] [doi: [10.7326/m19-3852](https://doi.org/10.7326/m19-3852)]

25. Liddy C, Drosinis P, Deri Armstrong C, McKellips F, Afkham A, Keely E. What are the cost savings associated with providing access to specialist care through the Champlain BASE eConsult service? A costing evaluation. *BMJ Open* 2016 Jun 23;6(6):e010920 [FREE Full text] [doi: [10.1136/bmjopen-2015-010920](https://doi.org/10.1136/bmjopen-2015-010920)] [Medline: [27338880](https://pubmed.ncbi.nlm.nih.gov/27338880/)]
26. Keely E, Liddy C, Afkham A. Utilization, benefits, and impact of an e-consultation service across diverse specialties and primary care providers. *Telemed J E Health* 2013 Oct;19(10):733-738 [FREE Full text] [doi: [10.1089/tmj.2013.0007](https://doi.org/10.1089/tmj.2013.0007)] [Medline: [23980939](https://pubmed.ncbi.nlm.nih.gov/23980939/)]
27. Dekhuijzen PNR, Smeele IJM, Smorenburg SM, Werkgroep Ketenzorg COPD. [Guideline for the non-pharmacological treatment of COPD]. *Ned Tijdschr Geneeskd* 2006 Jun 03;150(22):1233-1237. [Medline: [16796174](https://pubmed.ncbi.nlm.nih.gov/16796174/)]

Abbreviations

GP: general practitioner

ICPC: International Classification in Primary Care

Edited by A Kushniruk; submitted 21.09.21; peer-reviewed by A Tannoubi, AR Mohseni; comments to author 19.10.21; revised version received 11.11.21; accepted 20.01.22; published 01.04.22.

Please cite as:

Sanavro SM, van der Worp H, Jansen D, Koning P, Blanker MH, Prisma Platform Study Group

Evaluation of the First Year(s) of Physicians Collaboration on an Interdisciplinary Electronic Consultation Platform in the Netherlands:

Mixed Methods Observational Study

JMIR Hum Factors 2022;9(2):e33630

URL: <https://humanfactors.jmir.org/2022/2/e33630>

doi: [10.2196/33630](https://doi.org/10.2196/33630)

PMID: [35363155](https://pubmed.ncbi.nlm.nih.gov/35363155/)

©Sanne M Sanavro, Henk van der Worp, Danielle Jansen, Paul Koning, Marco H Blanker, Marco Prisma Platform Study Group. Originally published in *JMIR Human Factors* (<https://humanfactors.jmir.org>), 01.04.2022. This is an open-access article distributed under the terms of the Creative Commons Attribution License (<https://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in *JMIR Human Factors*, is properly cited. The complete bibliographic information, a link to the original publication on <https://humanfactors.jmir.org>, as well as this copyright and license information must be included.

Original Paper

Implementation of a Web-Based Tool for Shared Decision-making in Lung Cancer Screening: Mixed Methods Quality Improvement Evaluation

Julie Lowery^{1*}, PhD; Angela Fagerlin^{2,3}, PhD; Angela R Larkin^{1*}, BA; Renda S Wiener^{4,5}, MPH, MD; Sarah E Skurla^{1*}, MPH; Tanner J Caverly^{1,6,7*}, MPH, MD

¹Center for Clinical Management Research, Ann Arbor VA Healthcare System, Ann Arbor, MI, United States

²Department of Population Health Sciences, University of Utah School of Medicine, Salt Lake City, UT, United States

³Informatics Decision-Enhancement and Analytics Sciences Center for Innovation, VA Salt Lake City Healthcare System, Salt Lake City, MI, United States

⁴Center for Healthcare Organization & Implementation Research, VA Boston Healthcare System, Boston, MA, United States

⁵The Pulmonary Center, Boston University School of Medicine, Boston, MA, United States

⁶Department of Learning Health Sciences, University of Michigan School of Medicine, Ann Arbor, MI, United States

⁷Department of Internal Medicine, University of Michigan, Ann Arbor, MI, United States

*these authors contributed equally

Corresponding Author:

Tanner J Caverly, MPH, MD

Center for Clinical Management Research

Ann Arbor VA Healthcare System

2215 Fuller Road

Ann Arbor, MI, 48105

United States

Phone: 1 303 587 1038

Email: tcaverly@med.umich.edu

Abstract

Background: Lung cancer risk and life expectancy vary substantially across patients eligible for low-dose computed tomography lung cancer screening (LCS), which has important consequences for optimizing LCS decisions for different patients. To account for this heterogeneity during decision-making, web-based decision support tools are needed to enable quick calculations and streamline the process of obtaining individualized information that more accurately informs patient-clinician LCS discussions. We created DecisionPrecision, a clinician-facing web-based decision support tool, to help tailor the LCS discussion to a patient's individualized lung cancer risk and estimated net benefit.

Objective: The objective of our study is to test two strategies for implementing DecisionPrecision in primary care at eight Veterans Affairs medical centers: a quality improvement (QI) training approach and academic detailing (AD).

Methods: Phase 1 comprised a multisite, cluster randomized trial comparing the effectiveness of standard implementation (adding a link to DecisionPrecision in the electronic health record vs standard implementation plus the Learn, Engage, Act, and Process [LEAP] QI training program). The primary outcome measure was the use of DecisionPrecision at each site before versus after LEAP QI training. The second phase of the study examined the potential effectiveness of AD as an implementation strategy for DecisionPrecision at all 8 medical centers. Outcomes were assessed by comparing absolute tool use before and after AD visits and conducting semistructured interviews with a subset of primary care physicians (PCPs) following the AD visits.

Results: Phase 1 findings showed that sites that participated in the LEAP QI training program used DecisionPrecision significantly more often than the standard implementation sites (tool used 190.3, SD 174.8 times on average over 6 months at LEAP sites vs 3.5 SD 3.7 at standard sites; $P < .001$). However, this finding was confounded by the lack of screening coordinators at standard implementation sites. In phase 2, there was no difference in the 6-month tool use between before and after AD (95% CI -5.06 to 6.40; $P = .82$). Follow-up interviews with PCPs indicated that the AD strategy increased provider awareness and appreciation for the benefits of the tool. However, other priorities and limited time prevented PCPs from using them during routine clinical visits.

Conclusions: The phase 1 findings did not provide conclusive evidence of the benefit of a QI training approach for implementing a decision support tool for LCS among PCPs. In addition, phase 2 findings showed that our *light-touch*, single-visit AD strategy did not increase tool use. To enable tool use by PCPs, prediction-based tools must be fully automated and integrated into electronic health records, thereby helping providers personalize LCS discussions among their many competing demands. PCPs also need more time to engage in shared decision-making discussions with their patients.

Trial Registration: ClinicalTrials.gov NCT02765412; <https://clinicaltrials.gov/ct2/show/NCT02765412>

(*JMIR Hum Factors* 2022;9(2):e32399) doi:[10.2196/32399](https://doi.org/10.2196/32399)

KEYWORDS

shared decision-making; lung cancer; screening; clinical decision support; academic detailing; quality improvement; implementation

Introduction

Background

National lung cancer screening (LCS) guidelines have consistently recommended shared decision-making (SDM) to inform patients about the pros and cons of low-dose computed tomography (LDCT) screening and individualized LCS decisions [1,2]. The Centers for Medicare and Medicaid Services require documentation of SDM before initiating LDCT screening for its covered population, a policy that is unique among screening decisions in primary care [3]. Thus, understanding how to best implement and optimize SDM for LCS has been an urgent task for all health systems and clinicians offering LCS to their eligible patient population.

A key approach to SDM is to communicate accurate information about each person's potential to benefit from screening, especially if it meaningfully differs from the average summary results reported in trials. This is particularly important for LCS. Prior work examining the range of absolute benefits across all individuals enrolled in the National Lung Screening Trial demonstrated that conveying average population information can dramatically overstate or understate lung cancer mortality reduction in thousands of individuals [4]. This is because both lung cancer risk and life expectancy varied substantially across eligible patients, such that the average mortality benefit identified in the National Lung Screening Trial was driven upward by those at the highest risk, whereas many patients experienced a benefit that was far below the average [5,6]. Using prediction-based approaches to estimate individualized net benefits can support the communication of much more accurate information across individuals in a heterogeneous group of screening eligible individuals [7]. Such approaches form an inherently more patient-centered basis for SDM [8].

Objective

However, web-based decision tools that enable quick calculations and intuitive data presentations are needed to streamline the process of obtaining individualized information and more accurately inform patient-clinician LCS discussions in routine practice [9]. However, implementing clinical decision support tools in routine clinical practice has been difficult to achieve. Patient-facing tools have shown promise in improving patients' understanding of the criteria and procedural requirements for lung screening [10]; however, discussing the details of individualized risks and benefits with patients—a critical aspect of SDM—can be challenging for providers.

Numerous barriers, including infrastructure limitations and clinicians' perceptions of SDM taking too much time, have led to a lack of uptake in the integration of decision support interventions into standard clinical practice [11,12]. Therefore, the objective of our study is to test two strategies for implementing a prediction-based SDM tool for LCS (DecisionPrecision) [13] in primary care at eight Veterans Affairs (VA) medical centers: (1) a quality improvement (QI) training approach and (2) academic detailing (AD).

Methods

Overview

Our implementation efforts and evaluation of each took place in 2 phases. In phase 1, we used QI training as a strategy for implementing DecisionPrecision as part of a hybrid type 3 implementation study design [14]. Specifically, we used a multisite, cluster-based randomization trial to compare the effectiveness of standard implementation versus the effectiveness of standard implementation plus the Learn, Engage, Act, and Process (LEAP) QI training program [15]. The standard implementation comprised integrating a link to the tool into the VA computerized patient record system (CPRS) and providing educational materials on the tool to a local LCS champion.

Although we observed a substantial number of tool uses, primarily by dedicated screening coordinators (as described in the following sections), results from phase 1 suggested that neither LEAP nor standard implementation contributed to the absolute number of tool uses at a site by primary care providers (PCPs). Consequently, phase 2 of the study switched to a different implementation strategy—namely, AD—and the study design transitioned to a hybrid type 2 implementation study, which comprised coprimary aims: (1) to determine the effectiveness of the clinical intervention (ie, DecisionPrecision) on key outcomes and (2) to determine the potential effectiveness of AD as an implementation strategy for the intervention. In phase 2, the focus of this study was on the second coprimary aim. Findings from our evaluation of the effectiveness of the tool *in the LCS decision* (first coprimary aim) will be described elsewhere.

The primary outcome measure for both phases is the use of DecisionPrecision at the site level (ie, the absolute number of tool uses at a site over a specific period). This was the best measure of the reach of our implementation strategies that was feasible, given the constraints we faced: use of a standalone tool

that was deidentified and not connected to the health record and did not allow tracking of which clinicians and patients were associated with specific tool uses. This more ecologic site-level measure does not precisely fit the definition of reach in the reach, effectiveness, adoption, implementation, and maintenance (RE-AIM) framework [16], which defines reach as “the absolute number, proportion, and representativeness of individuals who are willing to participate in a given initiative, intervention, or program.” Moreover, although providers are the ones who must decide if they are willing to use the tool, it is the patients who we are ultimately trying to influence with tool use. Thus, we felt that it was more important to measure the number of tool uses rather than the number of providers using the tool. If 1 provider used the tool once and another provider used it 50 times, we were interested in the fact that the tool was used on 51 patients rather than the fact that it was used by 2 providers. A separate paper will focus on the effect of tool use on patient uptake of LCS (*effectiveness* from the RE-AIM framework).

In accordance with requirements of the journal, the CONSORT-EHEALTH (Consolidated Standards of Reporting Trials of Electronic and Mobile Health Applications and online Telehealth) checklist for reporting the study’s methods and findings was completed ([Multimedia Appendix 1](#)).

Common Methodology Across Phase 1 and Phase 2

Setting and Site Selection

We recruited 8 VA sites that participated in the Evaluation of the VA LCS Clinical Demonstration Project (LDCT demo) [17]. These sites were chosen as (1) they had an LCS program currently in place, and (2) as part of the LDCT demo, they used a standard set of clinical reminders built by the VA’s National Center for Health Promotion and Disease Prevention to identify patients eligible for LCS. Clinical reminders are embedded within the electronic health record of the Veterans Health Affairs (CPRS) and can prompt staff to screen patients, review or assess risk factors, or offer interventions and screenings that may be due for an individual patient. The use of these standardized reminders facilitates data collection on LCS eligibility, discussions, and decisions. These advantages provided an ideal situation for testing different implementation strategies for a decision support intervention.

Of these 8 sites, 7 (87%) agreed to participate in our implementation initiative, and 1 (13%) declined because of a competing program. To replace the eighth site, we used VA administrative data to identify sites that used the LCS clinical reminder in CPRS as an indicator of an active LCS program. One of these sites agreed to participate in the study.

Participants

To implement DecisionPrecision, we worked with clinical leaders from each site’s LCS program and screening coordinators when available. These core site team members enlisted others at their sites to help with implementation, including primary care leadership; PCPs; and team members from pulmonology, radiology, and oncology departments. Information technology and data security staff members were also engaged.

Intervention: DecisionPrecision

To meet the need for facilitating SDM and providing individualized information on LCS, we created DecisionPrecision, a provider-facing web-based decision support tool [13]. Our goal was to facilitate individualized and patient-centered SDM within the confines of a busy clinic environment. In particular, the tool seeks to quickly help PCPs tailor the LCS discussion based on the patient’s risk factor profile, more strongly encouraging screening for those with higher predicted lung cancer risk and potentially larger health gains with LCS. To this end, the tool provides the following: (1) individualized quantitative risk assessment of screening trade-offs, aided by an externally validated and accurate lung cancer risk prediction model [18], along with recommendation categories that clarify when screening is and is not preference sensitive [7]; (2) patient-friendly language for the provider to use with the patient; (3) patient-facing graphics, selected based on their ability to help patients understand personalized risks and benefits in a randomized survey experiment [19]; and (4) quick and easy documentation of personalized SDM after using the tool.

The final version of the DecisionPrecision tool tested in this implementation study was the result of multiple rounds of iterative changes that incorporated lessons from an observational field study of patient-clinician conversations in the absence of a decision support tool and iterative feedback and usability testing on multiple tool prototypes from decision aid researchers, PCPs, pulmonologists, screening coordinators, and patients. After phase 2, the tool also underwent extensive updates based on experiences and observations throughout the implementation project and additional feedback from clinicians, screening coordinators, and leadership. Screenshots of DecisionPrecision can be found in [Multimedia Appendix 2](#).

Ethics Approval

This project was categorized as QI and was, therefore, not formally reviewed by the institutional review boards of participating sites [20].

Phase 1: Evaluation of a QI Implementation Strategy

Phase 1 Methods

In phase 1, the 8 participating sites were randomized to either standard (S1, S2, S3, and S4) or enhanced implementation (ie, the LEAP QI training program: E1, E2, E3, and E4), stratified by the rate of clinician completion of the initial LCS clinical reminder (high vs low completion for patients eligible for LCS between May 2015 and June 2015). A CONSORT (Consolidated Standards of Reporting Trials) diagram summarizing the randomization of the sites is included in [Multimedia Appendix 3](#). The characteristics of the randomized sites are included in [Multimedia Appendix 4](#). Both standard and enhanced implementation strategies are described in detail in the following sections.

Standard Implementation

The standard implementation involved the following: (1) adding a tool link to the VA's electronic health record CPRS and (2) promoting the tool via emails, conference calls, and meetings.

DecisionPrecision Access Within the Electronic Health Record

A link to the decision tool was embedded within the clinical reminders for PCPs at all sites between August 29, 2017, and October 4, 2017. The language associated with the link and the specific location of the link were established based on conversations with the site team to best fit the mechanisms for the LCS at the site.

Promotion of DecisionPrecision

All sites were asked to notify relevant providers about DecisionPrecision. Site leads were given a draft email to providers that could be easily tailored with site-specific information. This email included a brief description of the key features of the tool and a link to a brief YouTube video that describes the tool's development (eg, how the algorithm was designed) and how to use the website. Educational materials on the tool were provided, including a sample risk assessment from the decision tool, a 1-page attachment explaining how to routinely and quickly use DecisionPrecision to personalize discussions about LCS, and a screenshot of the link in the CPRS clinical reminder. The site teams were asked to promote decision aid through key local leaders and staff meetings.

Enhanced Implementation

Enhanced implementation included all the components available in the standard implementation as well as QI training using the LEAP program. LEAP is a 26-week QI training program to engage frontline clinical teams in using a hands-on learning approach (see [Multimedia Appendix 5](#) for a brief description of the LEAP curriculum). The core components include a structured curriculum that focuses on QI methods, a coaching and learning community, and a web-based platform for sharing videos and other resources. Training goals include gaining confidence in applying QI methods to improve the quality of care within the demands of everyday clinical practice and completing a QI project using plan-do-study-act principles.

Of the 4 sites randomized to LEAP, 3 (75%) participated from February to July 2017 (E1, E2, and E4), and 1 (25%) declined (E3) because of time constraints for the site leads. The participating sites established an interdisciplinary LEAP team comprising 2 to 10 participants. The teams developed and executed a project charter to complete a plan-do-study-act cycle of change related to enhancing SDM for LCS using DecisionPrecision. The team members participating in LEAP were provided with early access to a link to DecisionPrecision in March 2017 so they could access the site as part of their improvement projects.

The improvement project at 1 site (E1) was the development of a process within 1 patient-aligned care team, whereby DecisionPrecision was used to identify eligible patients at the highest risk of lung cancer and then inform the SDM conversation for at least one veteran each week during the LEAP improvement program. The project for the other 2 sites (E2 and E4), which had more centralized screening programs and full-time screening coordinators, was for the screening coordinator to test the use of DecisionPrecision with all new consults for LCS.

Evaluation Methods

Phase 1 used a hybrid trial type 3 design [14]. The primary purpose of this design was to evaluate the effectiveness of the implementation strategy, and our primary question was, *Does enhanced implementation work better than standard implementation for facilitating the use of DecisionPrecision?* Thus, the primary outcome measure for evaluating the effectiveness of the enhanced implementation strategies was the use of DecisionPrecision at each participating site. The absolute tool use data by site were obtained from the DecisionPrecision website. In March 2018, a dropdown box was added asking the provider to indicate their site. To identify sites before March 2018, we used data on IP addresses collected by the website for each record entered and linked each IP address to a study site based on validation against the site data and IP addresses collected after March 2018. The analysis of these data included descriptive statistics of tool use by site. We also conducted brief interviews to assess the clinicians' impressions of the implementation strategy.

Phase 1 Results

As noted under the *Setting and Site Selection* section, sites were selected for this study based on their participation in the VA's LCS Clinical Demonstration Project to ensure that all sites had similar LCS programs. However, shortly after the start of the study, we observed that the sites had made some changes to their screening programs. Specifically, some sites had stopped the collection of data required to calculate lung cancer risk (specifically, detailed smoking histories and key data used by DecisionPrecision), some sites stopped the routine use of clinical reminders, some sites no longer used a screening coordinator for conducting SDM, and some sites no longer had a screening coordinator for performing any LCS tasks. As each of these changes had the potential to affect the use of DecisionPrecision, these changes across sites are summarized in [Table 1](#).

Consequently, our findings for this phase are presented by comparing the standard implementation sites to three different groups of facilities: (1) the original 4 facilities randomized to enhanced implementation (*intention to treat*), (2) the 3 facilities that participated in the enhanced implementation program (*as treated*), and (3) the 3 facilities that had a full-time screening coordinator engaged in SDM discussions with patients (*key resource scenario*).

Table 1. Summary of important changes occurring across the study sites after randomization.

Changes	Sites randomized to standard implementation				Sites randomized to enhanced implementation with LEAP ^a			
	S1	S2	S3	S4	E1	E2	E3 ^b	E4
Lung cancer screening clinical reminders (for providers)?	Limited	Yes	Limited	Yes	Limited	Yes	Yes	Yes
Screening coordinator for conducting shared decision-making?	No	No	No	No	No	Yes	Yes	Yes
Pack year reminder?	Limited	Yes	Limited	No	Yes	Yes	Yes	Yes

^aLEAP: Learn, Engage, Act, and Process.

^bAfter randomization, this site decided not to participate in LEAP but continued to participate in the overall trial.

Table 2 shows the data on tool use for all sites in the 6 months following the time at which they all had access to DecisionPrecision but before phase 2 (AD) was initiated in April 2018. It also indicates the sites in each comparison group.

Table 2. Tool use each month by site (number of patients).

Site	Enhanced ^a	LEAP ^{b,c}	Screening coordinator ^d	October 2017, n	November 2017, n	December 2017, n	January 2018, n	February 2018, n	March 2018, n	Total, n	Per month, mean (SD)
E1	✓	✓		7	1	5	2	4	4	23	3.8 (2.1)
E2	✓	✓	✓	54	80	52	74	72	85	417	69.5 (13.6)
E4	✓	✓	✓	42	47	40	40	30	34	233	38.8 (6.0)
E3	✓		✓	1	7	2	33	23	22	88	14.7 (13.2)
S1				2	0	0	0	0	0	2	0.3 (N/A ^e)
S2				1	0	0	0	0	0	1	0.2 (N/A)
S3				0	0	0	2	0	0	2	0.3 (N/A)
S4				3	2	1	2	0	1	9	1.5 (0.8)

^aRandomized to enhanced implementation.

^bLEAP: Learn, Engage, Act, and Process.

^cParticipated in LEAP QI training program.

^dStaffed by a lung cancer screening coordinator.

^eN/A: not applicable.

Table 3 presents the 6-month mean number of tool uses between the standard implementation sites versus the 3 comparison groups. All 3 comparisons showed significantly less tool use for the standard implementation sites. However, the presence of a screening coordinator at 3 of the enhanced implementation sites and none of the standard implementation sites greatly confounded these comparisons.

Table 3. Mean number of tool uses over 6 months: 4 standard implementation sites as compared with 3 ways of grouping intervention sites.

Comparisons	Tool uses: 6 months, mean (SD) ^a		Median difference <i>P</i> value ^b	Median difference (95% CI) ^b
	A sites	B sites		
Standard implementation sites (A) versus the original 4 facilities randomized to enhanced implementation (<i>intention to treat</i> ; B ₁)	3.5 (3.7)	190.3 (174.8)	.03	155.5 (14-416)
Standard implementation sites (A) versus the 3 facilities that participated in the enhanced implementation program (<i>as treated</i>); B ₂)	3.5 (3.7)	224.3 (197.1)	.049	231.0 (14-416)
Standard implementation sites (A) versus the 3 facilities that had a full-time screening coordinator engaged in SDM ^c discussions with patients (<i>key resource scenario</i> ; B ₃)	3.5 (3.7)	246.0 (164.9)	.049	231.0 (79-416)

^aTotal tool uses across all the sites in the group divided by the number of sites in the group.

^bOn the basis of the Wilcoxon rank-sum test using the medians of the differences (Hodges-Lehmann estimator) between sites in the 2 groups.

^cSDM: shared decision-making.

Phase 1 Discussion

Our phase 1 findings showed that sites that participated in the LEAP QI training program used DecisionPrecision significantly more often than the standard implementation sites. However, there is some indication that participation in LEAP was not the primary contributing reason for these findings, as the implementation arms were imbalanced with regard to the presence of a screening coordinator. As evidence, a site in the enhanced implementation group that did not participate in LEAP (E3) but had a screening coordinator adopted and used the tool. In addition, a site in the enhanced implementation group that did not adopt the tool in routine use (E1) was also the site in the group that no longer used a screening coordinator. Although we could not determine tool use by provider type (PCP vs screening coordinator) from the website data, the vast majority of tool uses at sites E2, E3, and E4 was by screening coordinators based on tool use data collected manually by the coordinators, which showed numbers comparable with those obtained from the website.

Although participation in the LEAP may have contributed to the increased use of the tool, we conclude that the existence of a screening coordinator likely played a much larger role in tool use. Therefore, the question remains whether the screening coordinators who adopted the tool would have used it to a lesser extent if they had not participated in LEAP. Although the site with a coordinator who did not participate in LEAP showed the lowest tool use of the 3 sites with coordinators, other data (not shown) showing tool use as a percentage of all eligible patients indicated that E3 had a comparable rate of tool use with that of E2 and E4.

In addition, feedback from one of the screening coordinators suggested that QI training was not key to implementing the tool:

Well, it [LEAP] seemed to be more tied towards quality process improvement, so it was helpful for that. When it comes to the decision precision tool, I'm not really sure if I can concretely tie it to that.

QI collaboratives, including internet-based videoconferencing adaptations, are a common approach to helping health care teams implement new initiatives or improve existing programs [21]. However, evidence for their success is mixed, including weaknesses in the reporting of methods and potential publication bias. Nevertheless, findings from several studies have shed light on some factors that are correlated with successful implementation. A study of 11 collaboratives focusing on 11 different topics found that innovation attributes, organizational support, innovative team culture, and professionals' commitment to change are instrumental to perceived effectiveness [22]. With specific regard to an innovation's attributes, the study found that the newer working methods were perceived by professionals as having relative benefit, being compatible with norms and values, not difficult to learn and implement, and leading to observable results, the more the implementation process was perceived as successful. This finding is certainly consistent with feedback from screening coordinators, who all perceived DecisionPrecision as doing an excellent job in conveying important information on risks and benefits to patients. They

found that the tool was relatively easy to use and incorporate into their workflow.

A systematic review of QI collaboratives concluded that collaboratives reporting success generally addressed relatively straightforward aspects of care and had a strong evidence base [21]. The implementation of DecisionPrecision could be considered straightforward in that it required only the addition of a link in the electronic medical record and did not require any significant changes in procedures or workflows. In addition, the scientific evidence underlying a prediction-based approach to LCS is relatively strong.

Furthermore, findings from a study on the effect of a learning collaborative on colorectal cancer screening rates in primary care practices are also consistent with ours [23]. Specifically, the teams had difficulty spreading the change beyond the clinicians who participated in the collaboration:

Other clinicians in a practice tended not to be aware of or engaged in the CRC (colorectal cancer) improvement efforts, and teams tended to communicate poorly with the rest of the practice regarding QI plans.

As a result—as occurred in our case—other providers, most notably PCPs, were not engaged and did not adopt the tool.

Another similarity between our study and the colorectal cancer screening study was that the clinicians who participated in the collaboratives (screening coordinators in our case) were very motivated to use the tool to improve the LCS process, which may not be the case among clinicians who did not participate (eg, PCPs).

Whether LEAP had a significant impact on the absolute number of DecisionPrecision uses by the screening coordinators, the collaborative approach did not have the intended effect of engaging the broader community of PCPs in using the tool. Feedback from participants in LEAP suggested the potential utility of an alternative strategy, namely one that focuses on one-on-one conversations with clinicians about the tool. Of the 26 sessions of the LEAP program, 1 (4%) was devoted to presenting and discussing the evidence behind the tool and how to use it. One screening coordinator noted the following:

I think the only helpful parts of it [LEAP], when it came to trying to implement the DecisionPrecision tool, was talking with the team...about what the stratified risks really mean...how you can come up with things like personalized harms and having that shared decision-making conversations where things are more preference-based—understanding that piece was extremely helpful and I can say that now, hindsight being 20/20 and having done a ton of shared decision-making in the last couple of years, I don't think I could've done it as effectively if I didn't have the knowledge that [was] shared with us during the LEAP program.

On the basis of the phase 1 findings, we switched to a different implementation strategy in phase 2—namely, AD.

AD was selected as an implementation strategy to convey information directly and one on one to PCPs about the evidence behind prediction-based screening and explain how to use DecisionPrecision. As noted on the National Resource for AD website [24], busy clinicians need an accurate source of current data on the effectiveness of current interventions. However, they have many competing demands for their time. Trying to assemble current evidence from a continuous influx of research is incredibly challenging to do on one's own. AD combines a one-on-one outreach approach with the best available evidence. We hired and trained a master's student in public health to meet clinicians to assess individual needs and then offer tailored, evidence-based advice for using DecisionPrecision as part of the LCS SDM process.

Phase 2: Academic Detailing

Phase 2 Methods

Our AD strategy focused on directly engaging PCPs, in addition to screening coordinators. We decided to use this strategy at all 8 participating sites rather than randomize sites to AD versus standard implementation in an effort to conduct a more extensive formative evaluation of the AD process, which, to the best of our knowledge, has not before been used to promote the use of a prediction-based SDM tool.

The goal of AD was to encourage providers, through the use of the DecisionPrecision tool, to adopt a prediction-based approach to tailoring how strongly screening is encouraged (based on estimated net benefit for the individual and consideration of how preference sensitive the decision is) to thereby facilitate a brief *everyday SDM* discussion and make decision-making more patient centered [7,8]. AD site visits were offered to all sites; of the 8 sites, 7 (87%) agreed to the site detailing visits. One of the sites underwent substantial workforce changes during the study and opted not to participate in the AD. Heads of primary care were asked to send emails to PCPs announcing the detailed visit and the purpose of the visit and encourage providers to participate in one-on-one detailing.

AD materials, which were developed and available for use during meetings, included a 4-page visual abstract of the evidence behind, benefits of, and key features of DecisionPrecision; a pocket card on how to use the tool; a CPRS clinical reminder screenshots and tool link handout; a handout on how to copy a templated description of the SDM discussion into the medical record; a list of references (in case of questions or concerns about the evidence); and a business card that included the URL to the DecisionPrecision website. The key information presented during these meetings were (1) how using a prediction-based approach for LCS can improve quality of care and (2) how to use the DecisionPrecision tool with eligible veterans to inform more patient-centered SDM and tailor screening encouragement during SDM discussions. At the end of each detailing session, the academic detailer asked for a provider's commitment to using the decision tool in the next 1 to 2 weeks and for permission to follow up with them 3 to 4 weeks after the detailing visit [25]. See [Multimedia Appendix](#)

6 for a summary of the characteristics of the AD strategy per the published recommendations.

We conducted semistructured phone interviews with a sample of PCPs for 2 to 4 weeks following their AD visits. The interviews included questions on the utility of the AD visit, use of DecisionPrecision since the visit, usefulness of the tool, ease of tool use, challenges in using the tool, and suggestions for improving the tool. Audio-recorded interviews lasted approximately 20 minutes and were transcribed verbatim.

Evaluation Methods

As noted in the *Introduction* section, this paper presents data on the potential effectiveness of AD as a strategy for promoting the use of DecisionPrecision. Data on the effectiveness of the tool as a clinical intervention for improving the quality of LCS decisions have been presented elsewhere. The effectiveness of AD as an implementation strategy was assessed by (1) examining tool use following the AD visits and (2) conducting semistructured interviews with a subset of PCPs following their participation in AD visits.

Analysis of Tool Use

We conducted an interrupted time-series analysis to determine whether there was a difference in the overall tool use between the 6 months following the initiation of AD and the 6 months following the initiation of enhanced implementation (and before AD). We fitted a linear mixed model with the study period as the fixed effect of interest and a random intercept for each site.

Post-AD Interviews

We used NVivo (version 12; QSR International) to conduct an inductive thematic content analysis of the postdetailing interviews, searching for themes that emerged from the qualitative data. Team discussion of the findings led to agreement on the common themes, which included the major barriers to tool use and the features of the tool that the providers found to be beneficial.

Phase 2 Results

We examined 105 PCPs at the 7 participating sites from June to October 2018 (E3 chose not to participate). Each site visit lasted 2 to 3 days, except for E1, where visits occurred over 2 months. The academic detailer met providers in primary care clinics, primary care resident clinics, and community-based outpatient clinics. Snowball sampling was used to identify providers before and during the site visit. Individual meetings were tailored to provider needs in terms of both content and duration. The duration of the meetings ranged from 4 to 40 minutes, with a mean duration of 13 minutes. Most meetings were one on one; however, a few meetings were with 2 to 3 providers simultaneously.

Tool Use by Site

[Table 4](#) shows data on absolute tool use for the 6 months following the initiation of AD for all sites participating in AD. For comparison purposes, the last 2 columns of the table also show the total and average tool use for a similar period (6 months) before the AD intervention.

Table 4. Monthly tool use at seven sites participating in AD^{a,b} (number of patients).

Sites	Before AD						Total 6 months after AD	Per month after AD, mean (SD)	Total 6 months before AD ^c	Per month before AD, ^c mean (SD)
	October 2018, n	November 2018, n	December 2018, n	January 2019, n	February 2019, n	March 2019, n				
All sites	84	90	114	137	137	153	715	119.2 (27.9)	687	114.5 (12.1)
E1	1	4	2	2	4	1	14	2.3 (1.4)	23	3.8 (2.1)
E2	52	62	65	72	73	94	418	69.7 (14.2)	417	69.5 (13.6)
E4	20	20	45	51	51	37	224	37.3 (14.4)	233	38.8 (6.0)
S1	4	4	1	5	2	7	23	3.8 (2.1)	2	0.3 (N/A ^d)
S2	0	0	0	1	0	1	2	0.3 (N/A)	1	0.2 (N/A)
S3	7	0	1	4	7	13	32	5.3 (4.4)	2	0.3 (N/A)
S4	0	0	0	2	0	0	2	0.3 (N/A)	9	1.5 (0.8)

^aAD: academic detailing.

^bThe site designations, Enhanced Implementation and Standard Implementation, are not relevant for phase 2 as all sites received academic detailing; however, the labeling was maintained for linking to the phase 1 data (last 2 columns).

^cPre-AD months: October 2017 to March 2018; data are pulled from [Table 2](#).

^dN/A: not available.

An interrupted time-series analysis showed no significant difference in tool use between pre- and post-AD periods (95% CI 5.06-6.40, fewer tool uses after AD to more tool uses after AD; $P=.82$). Thus, it appears that the introduction of AD as an implementation strategy did not encourage substantial additional tool use (see [Multimedia Appendix 7](#) for detailed results of this analysis).

Interview Responses

Of the 105 PCPs who participated in the AD, 83 (79%) provided their contact information for follow-up purposes. Of the 83 providers contacted, 33 (40%) participated in the post-AD follow-up interviews. Virtually unanimously, the participants felt that the AD visit helped provide them with an explanation of the tool's purpose, the science underlying its development, and how to use it. Respondents appreciated the concise presentation and opportunity for a one-on-one discussion to walk them through the tool. Regarding the tool itself, many saw value in the tool's ability to (1) shape clinician feelings about the LCS and (2) guide a useful approach to LCS discussions. They also felt that the tool enhanced their ability to share information about individualized risks and the pros and cons of screening. However, at follow-up, a few PCPs (6/33, 18%) used the tool with an actual patient. Limited time in the clinic was perceived as a key barrier by almost all the PCPs. Most PCPs reported needing 1 to 2 minutes to discuss LCS but frequently voiced not having even 1 to 2 minutes during a visit because of patient-specific needs that were a higher priority (eg, acute complaints) or organizational priorities (eg, performance measures). Similarly, having to input clinical data on risk factors into the tool was seen as a significant barrier to tool use as it added more time to the visit.

Phase 2 Discussion

In phase 2, there was no difference in tool use between before and after AD. Follow-up interviews with PCPs indicated that the AD strategy increased provider awareness and appreciation for the benefits of the tool. However, other priorities and limited time prevented PCPs from using them during routine clinical visits.

We decided to pursue AD as a strategy because of the consistent literature documenting the effectiveness of this approach for aligning clinician behavior with evidence-based best practices [26-29]. Others have emphasized that for eHealth to optimize preventive care, electronic risk factor data need to be seen as relevant and useful by PCPs [30]. Before our AD intervention, we had limited success in engaging PCPs in considering the important role of overall lung cancer when making LCS decisions. We needed a strategy that could help us have meaningful conversations with frontline clinicians making daily decisions about LCS.

Although rooted in a strong evidence base, prediction-based approaches to decision-making about cancer screening are relatively novel and unfamiliar to many clinicians [31,32]. There are potential cognitive challenges with moving from SDM that conveys population average information to a one-size-fits-all approach for all patients who meet eligibility criteria and toward a prediction-based approach that tailors the strength of the recommendation based on the degree of estimated net benefit for each eligible individual. Given the frequency of cancer screening decisions in primary care and how entrenched practice styles and decision-making can be for such practices, it was unclear whether a brief AD intervention would be able to successfully convey the rationale for a prediction-based approach. Most AD studies have focused on medication prescribing, and few such studies used detailing to modify a

decision-making approach to a commonly delivered category of service such as cancer screening [29].

On the basis of feedback from follow-up interviews with participating providers, AD allowed us to meaningfully engage dozens of PCPs across 7 sites. At the very least, these providers are now familiar with the prediction-based approach to SDM for LCS, the tool, and how it can be used in a busy primary care setting. Interview responses suggest that we have changed the way some clinicians think about decision-making for LCS, especially their understanding of the utility of a prediction-based approach to screening decisions. Nevertheless, our light-touch AD did not result in the routine use of the tool among PCPs in our study sample, primarily because of time constraints. As others have noted, the incentives (and disincentives) in our health care system will need to change if providers are to have sufficient time to engage in SDM [33].

General Discussion

Limitations

The limitations of this study include the small sample size (8 sites). Other studies have emphasized the importance of examining whether specific components of multicomponent implementation strategies have stronger associations with absolute tool use than other components in an effort to streamline these strategies to be more time and cost-effective [34]. A larger sample size would have enabled us to examine whether specific components of LEAP were associated with the use of DecisionPrecision. An additional limitation was that enhanced implementation was confounded by the presence of a screening coordinator, making it impossible to attribute the findings to the enhanced implementation strategy alone. In phase 2, the AD intervention was based on a single site visit and a single one-on-one conversation with PCPs rather than multiple reinforcing visits or a longitudinal relationship with PCPs.

Conclusions

The phase 1 findings do not provide conclusive evidence of the benefit of a QI training approach for implementing a decision support tool for LCS among PCPs. Screening coordinators in the study used the tool frequently, and it is possible that the LEAP program helped them adopt the tool. However, other factors may have contributed to tool use, including the coordinators' perception of the added benefit of using the tool

as part of their responsibility for educating patients on LCS and the relative ease of incorporating tool use into their workflow.

As PCPs were not engaged in the phase 1 implementation strategy, the phase 2 implementation strategy—AD—targeted these clinicians. On the basis of our experience with phase 1, the focus of the AD approach was to educate PCPs on the benefits of tool use and discuss the best ways of incorporating it into their clinical practice. However, even when PCPs see value in a prediction-based approach to LCS decision-making and a tool to support that approach, they face major challenges in implementing it in a busy primary care clinic. This was a consistent finding across all study sites. Thus, in terms of the RE-AIM framework, we feel that the adoption (willingness to initiate a program) and reach (willingness to use) of our standalone DecisionPrecision tool, if left unchanged, is likely to be limited among PCPs within other health care settings as well.

One implication of these findings for implementing decision support tools for LCS—and potentially other cancer screening tools—is that QI as an implementation strategy may not be helpful; instead, the focus of implementation should be on working with individual clinicians and screening coordinators to promote tool use. Screening coordinators bought into the rationale for using the tool *and* were able to adopt and use it routinely (high awareness and good reach among coordinators). However, our *light-touch*, single-visit AD strategy did not affect tool use among PCPs, although feedback from PCPs suggested that this strategy did achieve our goal of increasing provider awareness and appreciation of the benefits of the tool (awareness but limited adoption and reach among PCPs). Other priorities and limited time prevented PCPs from using them during routine primary care clinic visits. These barriers point to the second implication of our findings; namely, prediction-based SDM tools need to be automated as much as possible for use in primary care to better integrate into workflows and help PCPs more quickly understand how to prioritize LCS discussions among other competing demands [35]. Regarding the latter need, an ongoing Agency for Health Research and Quality-funded project is addressing this barrier by automating predictions and integrating the DecisionPrecision tool within multiple electronic health records, including Epic, Cerner, and CPRS health record systems. CPRS is still used in most VA health systems at present before a planned transition to Cerner [36-39].

Acknowledgments

TJC is supported by a Career Development Award from Veterans Affairs Health Services Research and Development (CDA 16-151). This study was conducted with a grant from the Veterans Affairs Quality Enhancement Research Initiative Program (QUE 15-286).

Conflicts of Interest

At the time of the study, all the authors were employees of the Department of Veterans Affairs (VA). This work was funded by the Department of VA Health Services Research and Development Quality Enhancement Research Initiative grant (QUE 15-286) and the VA Career Development Award (CDA 16-151). TJC and AF created an invention and submitted an invention disclosure form with the VA for the web-based lung cancer screening (LCS) risk calculator, screenLC (previously known as, and referenced throughout the manuscript, as DecisionPrecision) [19], but did not receive financial support for this invention. RSW participates

as a panelist and coauthors several guidelines or papers related to the implementation of LCS and serves as a co-chair of the VISN1 LCS implementation task force but does not receive dedicated salary support for this role. JL is the chair of VA's Scientific Merit Review Board, which awards VA Health Services Research and Development grant funds. This organization funded the project on which we are reporting but was not involved in any way in the review and discussions regarding this grant. Authors TJC, AF, RSW, and JL also coauthored a related manuscript [40].

Multimedia Appendix 1

CONSORT-EHEALTH checklist (V 1.6.1).

[PDF File (Adobe PDF File), 658 KB - [humanfactors_v9i2e32399_app1.pdf](#)]

Multimedia Appendix 2

DecisionPrecision screenshots.

[DOCX File , 340 KB - [humanfactors_v9i2e32399_app2.docx](#)]

Multimedia Appendix 3

CONSORT (Consolidated Standards of Reporting Trials) diagram, phase 1.

[DOC File , 32 KB - [humanfactors_v9i2e32399_app3.doc](#)]

Multimedia Appendix 4

Site characteristics.

[DOCX File , 17 KB - [humanfactors_v9i2e32399_app4.docx](#)]

Multimedia Appendix 5

Learn, Engage, Act, and Process (LEAP) curriculum.

[PNG File , 213 KB - [humanfactors_v9i2e32399_app5.png](#)]

Multimedia Appendix 6

Characteristics of the academic detailing strategy.

[DOCX File , 15 KB - [humanfactors_v9i2e32399_app6.docx](#)]

Multimedia Appendix 7

Results of the interrupted time-series analysis.

[DOCX File , 14 KB - [humanfactors_v9i2e32399_app7.docx](#)]

References

1. Mazzone PJ, Silvestri GA, Patel S, Kanne JP, Kinsinger LS, Wiener RS, et al. Screening for lung cancer: CHEST guideline and expert panel report. *Chest* 2018;153(4):954-985. [doi: [10.1016/j.chest.2018.01.016](#)] [Medline: [29374513](#)]
2. US Preventive Services Task Force, Krist AH, Davidson KW, Mangione CM, Barry MJ, Cabana M, et al. Screening for lung cancer: US preventive services task force recommendation statement. *JAMA* 2021;325(10):962-970. [doi: [10.1001/jama.2021.1117](#)] [Medline: [33687470](#)]
3. Hughes DL. CMS proposes coverage for lung cancer screening with low dose CT. *Health Affairs Blog*. 2014. URL: <http://healthaffairs.org/blog/2014/12/09/cms-proposes-coverage-with-evidence-development-for-lung-cancer-screening-with-low-dose-ct/> [accessed 2014-12-29]
4. Kovalchik SA, Tammemagi M, Berg CD, Caporaso NE, Riley TL, Korch M, et al. Targeting of low-dose CT screening according to the risk of lung-cancer death. *N Engl J Med* 2013;369(3):245-254 [FREE Full text] [doi: [10.1056/NEJMoa1301851](#)] [Medline: [23863051](#)]
5. Bach PB, Gould MK. When the average applies to no one: personalized decision making about potential benefits of lung cancer screening. *Ann Intern Med* 2012;157(8):571-573. [doi: [10.7326/0003-4819-157-8-201210160-00524](#)] [Medline: [22893040](#)]
6. Caverly TJ, Meza R. Using risk models to make lung cancer screening decisions: evidence-based and getting better. *Ann Intern Med* 2019;171(9):669-670. [doi: [10.7326/M19-2869](#)] [Medline: [31634912](#)]
7. Caverly TJ, Cao P, Hayward RA, Meza R. Identifying patients for whom lung cancer screening is preference-sensitive: a microsimulation study. *Ann Intern Med* 2018;169(1):1-9 [FREE Full text] [doi: [10.7326/M17-2561](#)] [Medline: [29809244](#)]
8. Caverly TJ, Hayward RA. Dealing with the lack of time for detailed shared decision-making in primary care: everyday shared decision-making. *J Gen Intern Med* 2020;35(10):3045-3049 [FREE Full text] [doi: [10.1007/s11606-020-06043-2](#)] [Medline: [32779137](#)]

9. Van de Velde S, Kunnamo I, Roshanov P, Kortteisto T, Aertgeerts B, Vandvik PO, GUIDES expert panel. The GUIDES checklist: development of a tool to improve the successful use of guideline-based computerised clinical decision support. *Implement Sci* 2018;13(1):86 [FREE Full text] [doi: [10.1186/s13012-018-0772-3](https://doi.org/10.1186/s13012-018-0772-3)] [Medline: [29941007](https://pubmed.ncbi.nlm.nih.gov/29941007/)]
10. Carter-Harris L, Comer RS, Slaven Ii JE, Monahan PO, Vode E, Hanna NH, et al. Computer-tailored decision support tool for lung cancer screening: community-based pilot randomized controlled trial. *J Med Internet Res* 2020;22(11):e17050 [FREE Full text] [doi: [10.2196/17050](https://doi.org/10.2196/17050)] [Medline: [33141096](https://pubmed.ncbi.nlm.nih.gov/33141096/)]
11. Friedberg MW, Van Busum K, Wexler R, Bowen M, Schneider EC. A demonstration of shared decision making in primary care highlights barriers to adoption and potential remedies. *Health Aff (Millwood)* 2013;32(2):268-275. [doi: [10.1377/hlthaff.2012.1084](https://doi.org/10.1377/hlthaff.2012.1084)] [Medline: [23381519](https://pubmed.ncbi.nlm.nih.gov/23381519/)]
12. Elwyn G, Scholl I, Tietbohl C, Mann M, Edwards AG, Clay C, et al. "Many miles to go ...": a systematic review of the implementation of patient decision support interventions into routine clinical practice. *BMC Med Inform Decis Mak* 2013;13 Suppl 2:S14 [FREE Full text] [doi: [10.1186/1472-6947-13-S2-S14](https://doi.org/10.1186/1472-6947-13-S2-S14)] [Medline: [24625083](https://pubmed.ncbi.nlm.nih.gov/24625083/)]
13. Caverly TJ, Fagerlin A. Lung cancer screening risk calculator. ScreenLC. URL: <https://screenlc.com/> [accessed 2020-07-22]
14. Curran GM, Bauer M, Mittman B, Pyne JM, Stetler C. Effectiveness-implementation hybrid designs: combining elements of clinical effectiveness and implementation research to enhance public health impact. *Med Care* 2012;50(3):217-226 [FREE Full text] [doi: [10.1097/MLR.0b013e3182408812](https://doi.org/10.1097/MLR.0b013e3182408812)] [Medline: [22310560](https://pubmed.ncbi.nlm.nih.gov/22310560/)]
15. Damschroder LJ, Yankey NR, Robinson CH, Freitag MB, Burns JA, Raffa SD, et al. The LEAP program: quality improvement training to address team readiness gaps identified by implementation science findings. *J Gen Intern Med* 2021;36(2):288-295 [FREE Full text] [doi: [10.1007/s11606-020-06133-1](https://doi.org/10.1007/s11606-020-06133-1)] [Medline: [32901440](https://pubmed.ncbi.nlm.nih.gov/32901440/)]
16. Glasgow RE, Harden SM, Gaglio B, Rabin B, Smith ML, Porter GC, et al. RE-AIM planning and evaluation framework: adapting to new science and practice with a 20-year review. *Front Public Health* 2019;7:64 [FREE Full text] [doi: [10.3389/fpubh.2019.00064](https://doi.org/10.3389/fpubh.2019.00064)] [Medline: [30984733](https://pubmed.ncbi.nlm.nih.gov/30984733/)]
17. Kinsinger LS, Anderson C, Kim J, Larson M, Chan SH, King HA, et al. Implementation of lung cancer screening in the veterans health administration. *JAMA Intern Med* 2017;177(3):399-406. [doi: [10.1001/jamainternmed.2016.9022](https://doi.org/10.1001/jamainternmed.2016.9022)] [Medline: [28135352](https://pubmed.ncbi.nlm.nih.gov/28135352/)]
18. Bach PB, Kattan MW, Thornquist MD, Kris MG, Tate RC, Barnett MJ, et al. Variations in lung cancer risk among smokers. *J Natl Cancer Inst* 2003;95(6):470-478. [doi: [10.1093/jnci/95.6.470](https://doi.org/10.1093/jnci/95.6.470)] [Medline: [12644540](https://pubmed.ncbi.nlm.nih.gov/12644540/)]
19. Bonner C, Trevena LJ, Gaissmaier W, Han PK, Okan Y, Ozanne E, et al. Current best practice for presenting probabilities in patient decision aids: fundamental principles. *Med Decis Making* 2021;41(7):821-833. [doi: [10.1177/0272989X211996328](https://doi.org/10.1177/0272989X211996328)] [Medline: [33660551](https://pubmed.ncbi.nlm.nih.gov/33660551/)]
20. Program guide for quality improvement projects. VA Ann Arbor Healthcare System. 2021. URL: https://www.annarbor.research.va.gov/ANNARBORRESEARCH/docs/RDC/GUIDE_QI.pdf [accessed 2022-03-25]
21. Wells S, Tamir O, Gray J, Naidoo D, Bekhit M, Goldmann D. Are quality improvement collaboratives effective? A systematic review. *BMJ Qual Saf* 2018;27(3):226-240. [doi: [10.1136/bmjqs-2017-006926](https://doi.org/10.1136/bmjqs-2017-006926)] [Medline: [29055899](https://pubmed.ncbi.nlm.nih.gov/29055899/)]
22. Strating MM, Nieboer AP. Explaining variation in perceived team effectiveness: results from eleven quality improvement collaboratives. *J Clin Nurs* 2013;22(11-12):1692-1706. [doi: [10.1111/j.1365-2702.2012.04120.x](https://doi.org/10.1111/j.1365-2702.2012.04120.x)] [Medline: [22612406](https://pubmed.ncbi.nlm.nih.gov/22612406/)]
23. Shaw EK, Ohman-Strickland PA, Piasecki A, Hudson SV, Ferrante JM, McDaniel RR, et al. Effects of facilitated team meetings and learning collaboratives on colorectal cancer screening rates in primary care practices: a cluster randomized trial. *Ann Fam Med* 2013;11(3):220-228 [FREE Full text] [doi: [10.1370/afm.1505](https://doi.org/10.1370/afm.1505)] [Medline: [23690321](https://pubmed.ncbi.nlm.nih.gov/23690321/)]
24. Improving care via 1:1 dialogue. National Resource Center for Academic Detailing [NaRCAD]. URL: <https://www.narcad.org/about-us.html> [accessed 2021-07-22]
25. Van Hoof TJ, Harrison LG, Miller NE, Pappas MS, Fischer MA. Characteristics of academic detailing: results of a literature review. *Am Health Drug Benefits* 2015;8(8):414-422 [FREE Full text] [Medline: [26702333](https://pubmed.ncbi.nlm.nih.gov/26702333/)]
26. Avorn J, Fischer M. 'Bench to behavior': translating comparative effectiveness research into improved clinical practice. *Health Aff (Millwood)* 2010;29(10):1891-1900. [doi: [10.1377/hlthaff.2010.0696](https://doi.org/10.1377/hlthaff.2010.0696)] [Medline: [20921491](https://pubmed.ncbi.nlm.nih.gov/20921491/)]
27. Grimshaw JM, Thomas RE, MacLennan G, Fraser C, Ramsay C, Vale L, et al. Effectiveness and efficiency of guideline dissemination and implementation strategies. *Int J Technol Assess Health Care* 2005;21(1):149. [doi: [10.1017/s0266462305290190](https://doi.org/10.1017/s0266462305290190)]
28. Avorn J, Soumerai SB. Improving drug-therapy decisions through educational outreach. A randomized controlled trial of academically based "detailing". *N Engl J Med* 1983;308(24):1457-1463. [doi: [10.1056/NEJM198306163082406](https://doi.org/10.1056/NEJM198306163082406)] [Medline: [6406886](https://pubmed.ncbi.nlm.nih.gov/6406886/)]
29. O'Brien MA, Rogers S, Jamtvedt G, Oxman AD, Odgaard-Jensen J, Kristoffersen DT, et al. Educational outreach visits: effects on professional practice and health care outcomes. *Cochrane Database Syst Rev* 2007(4):CD000409 [FREE Full text] [doi: [10.1002/14651858.CD000409.pub2](https://doi.org/10.1002/14651858.CD000409.pub2)] [Medline: [17943742](https://pubmed.ncbi.nlm.nih.gov/17943742/)]
30. Carey M, Noble N, Mansfield E, Waller A, Henskens F, Sanson-Fisher R. The role of eHealth in optimizing preventive care in the primary care setting. *J Med Internet Res* 2015;17(5):e126 [FREE Full text] [doi: [10.2196/jmir.3817](https://doi.org/10.2196/jmir.3817)] [Medline: [26001983](https://pubmed.ncbi.nlm.nih.gov/26001983/)]

31. Kent DM, van Klaveren D, Paulus JK, D'Agostino R, Goodman S, Hayward R, et al. The predictive approaches to treatment effect heterogeneity (PATH) statement: explanation and elaboration. *Ann Intern Med* 2020;172(1):W1-25 [FREE Full text] [doi: [10.7326/M18-3668](https://doi.org/10.7326/M18-3668)] [Medline: [31711094](https://pubmed.ncbi.nlm.nih.gov/31711094/)]
32. Kent DM, Steyerberg E, van Klaveren D. Personalized evidence based medicine: predictive approaches to heterogeneous treatment effects. *BMJ* 2018;363:k4245 [FREE Full text] [doi: [10.1136/bmj.k4245](https://doi.org/10.1136/bmj.k4245)] [Medline: [30530757](https://pubmed.ncbi.nlm.nih.gov/30530757/)]
33. Jimbo M, Shultz CG, Nease DE, Fetters MD, Power D, Ruffin 4th MT. Perceived barriers and facilitators of using a web-based interactive decision aid for colorectal cancer screening in community practice settings: findings from focus groups with primary care clinicians and medical office staff. *J Med Internet Res* 2013;15(12):e286 [FREE Full text] [doi: [10.2196/jmir.2914](https://doi.org/10.2196/jmir.2914)] [Medline: [24351420](https://pubmed.ncbi.nlm.nih.gov/24351420/)]
34. Walunas TL, Ye J, Bannon J, Wang A, Kho AN, Smith JD, et al. Does coaching matter? Examining the impact of specific practice facilitation strategies on implementation of quality improvement interventions in the Healthy Hearts in the Heartland study. *Implement Sci* 2021;16(1):33 [FREE Full text] [doi: [10.1186/s13012-021-01100-8](https://doi.org/10.1186/s13012-021-01100-8)] [Medline: [33789696](https://pubmed.ncbi.nlm.nih.gov/33789696/)]
35. Kawamoto K, Houlihan CA, Balas EA, Lobach DF. Improving clinical practice using clinical decision support systems: a systematic review of trials to identify features critical to success. *BMJ* 2005;330(7494):765 [FREE Full text] [doi: [10.1136/bmj.38398.500764.8F](https://doi.org/10.1136/bmj.38398.500764.8F)] [Medline: [15767266](https://pubmed.ncbi.nlm.nih.gov/15767266/)]
36. Scalable decision support and shared decisionmaking for lung cancer screening. Agency for Healthcare Research and Quality, Digital Healthcare Research, Informing Improvement in Care Quality, Safety, and Efficiency. URL: <https://digital.ahrq.gov/ahrq-funded-projects/scalable-decision-support-and-shared-decisionmaking-lung-cancer-screening> [accessed 2021-04-01]
37. Kawamoto K, Kukhareva PV, Weir C, Flynn MC, Nanjo CJ, Martin DK, et al. Establishing a multidisciplinary initiative for interoperable electronic health record innovations at an academic medical center. *JAMIA Open* 2021;4(3):ooab041 [FREE Full text] [doi: [10.1093/jamiaopen/ooab041](https://doi.org/10.1093/jamiaopen/ooab041)] [Medline: [34345802](https://pubmed.ncbi.nlm.nih.gov/34345802/)]
38. Reese TJ, Schlechter CR, Kramer H, Kukhareva P, Weir CR, Del Fiol G, et al. Implementing lung cancer screening in primary care: needs assessment and implementation strategy design. *Transl Behav Med* 2022;12(2):187-197. [doi: [10.1093/tbm/ibab115](https://doi.org/10.1093/tbm/ibab115)] [Medline: [34424342](https://pubmed.ncbi.nlm.nih.gov/34424342/)]
39. Kawamoto K. Decision precision+: increasing lung cancer screening for at-risk patients. Agency for Healthcare Research and Quality. 2020. URL: <https://digital.ahrq.gov/2019-year-review/research-summary/emerging-innovative-newly-funded-research/decision-precision-increasing-lung-cancer-screening-risk-patients> [accessed 2021-09-16]
40. Leishman NJ, Wiener RS, Fagerlin A, Hayward RA, Lowery J, Caverly TJ. Variation in eligible patients' agreeing to and receiving lung cancer screening: a cohort study. *Am J Prev Med* 2021;60(4):520-528. [doi: [10.1016/j.amepre.2020.10.014](https://doi.org/10.1016/j.amepre.2020.10.014)] [Medline: [33342671](https://pubmed.ncbi.nlm.nih.gov/33342671/)]

Abbreviations

AD: academic detailing

CONSORT: Consolidated Standards of Reporting Trials

CONSORT-EHEALTH: Consolidated Standards of Reporting Trials of Electronic and Mobile Health Applications and online Telehealth

CPRS: computerized patient record system

LCS: lung cancer screening

LDCT: low-dose computed tomography

LEAP: Learn, Engage, Act, and Process

PCP: primary care physician

QI: quality improvement

RE-AIM: reach, effectiveness, adoption, implementation, and maintenance

SDM: shared decision-making

VA: Veterans Affairs

Edited by A Kushniruk; submitted 26.07.21; peer-reviewed by G Taksler, J Ye; comments to author 06.09.21; revised version received 28.10.21; accepted 28.11.21; published 01.04.22.

Please cite as:

Lowery J, Fagerlin A, Larkin AR, Wiener RS, Skurla SE, Caverly TJ

Implementation of a Web-Based Tool for Shared Decision-making in Lung Cancer Screening: Mixed Methods Quality Improvement Evaluation

JMIR Hum Factors 2022;9(2):e32399

URL: <https://humanfactors.jmir.org/2022/2/e32399>

doi: [10.2196/32399](https://doi.org/10.2196/32399)

PMID: [35363144](https://pubmed.ncbi.nlm.nih.gov/35363144/)

©Julie Lowery, Angela Fagerlin, Angela R Larkin, Renda S Wiener, Sarah E Skurla, Tanner J Caverly. Originally published in JMIR Human Factors (<https://humanfactors.jmir.org>), 01.04.2022. This is an open-access article distributed under the terms of the Creative Commons Attribution License (<https://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Human Factors, is properly cited. The complete bibliographic information, a link to the original publication on <https://humanfactors.jmir.org>, as well as this copyright and license information must be included.

Original Paper

Requirements for a Bespoke Intensive Care Unit Dashboard in Response to the COVID-19 Pandemic: Semistructured Interview Study

Brittany Davidson^{1,2}, MA, PhD; Katuska Mara Ferrer Portillo³, LLB, MA, MSc; Marceli Wac^{1,4}, BSc; Chris McWilliams^{1,4}, BSc, MSc, PhD; Christopher Bourdeaux⁴, BCHIR, MA, MB; Ian Craddock¹, BEng, PhD

¹Department of Electrical & Electronic Engineering, University of Bristol, Bristol, United Kingdom

²School of Management, University of Bath, Bath, United Kingdom

³French Department, School of Modern Languages, University of Bristol, Bristol, United Kingdom

⁴University Hospitals Bristol and Weston National Health Service Foundation Trust, Bristol, United Kingdom

Corresponding Author:

Marceli Wac, BSc

Department of Electrical & Electronic Engineering

University of Bristol

1 Cathedral Square

Bristol, BS1 5DD

United Kingdom

Phone: 44 07774286342

Email: m.wac@bristol.ac.uk

Abstract

Background: Intensive care units (ICUs) around the world are in high demand due to patients with COVID-19 requiring hospitalization. As researchers at the University of Bristol, we were approached to develop a bespoke data visualization dashboard to assist two local ICUs during the pandemic that will centralize disparate data sources in the ICU to help reduce the cognitive load on busy ICU staff in the ever-evolving pandemic.

Objective: The aim of this study was to conduct interviews with ICU staff in University Hospitals Bristol and Weston National Health Service Foundation Trust to elicit requirements for a bespoke dashboard to monitor the high volume of patients, particularly during the COVID-19 pandemic.

Methods: We conducted six semistructured interviews with clinical staff to obtain an overview of their requirements for the dashboard and to ensure its ultimate suitability for end users. Interview questions aimed to understand the job roles undertaken in the ICU, potential uses of the dashboard, specific issues associated with managing COVID-19 patients, key data of interest, and any concerns about the introduction of a dashboard into the ICU.

Results: From our interviews, we found the following design requirements: (1) *a flexible dashboard*, where the functionality can be updated quickly and effectively to respond to emerging information about the management of this new disease; (2) *a mobile dashboard*, which allows staff to move around on wards with a dashboard, thus potentially replacing paper forms to enable detailed and consistent data entry; (3) *a customizable and intuitive dashboard*, where individual users would be able to customize the appearance of the dashboard to suit their role; (4) *real-time data and trend analysis* via informative data visualizations that help busy ICU staff to understand a patient's clinical trajectory; and (5) *the ability to manage tasks and staff*, tracking both staff and patient movements, handovers, and task monitoring to ensure the highest quality of care.

Conclusions: The findings of this study confirm that digital solutions for ICU use would potentially reduce the cognitive load of ICU staff and reduce clinical errors at a time of notably high demand of intensive health care.

(*JMIR Hum Factors* 2022;9(2):e30523) doi:[10.2196/30523](https://doi.org/10.2196/30523)

KEYWORDS

intensive care; critical care; COVID-19; human-centered design; dashboard; eHealth; disease monitoring; monitoring; ICU; design; development; interview

Introduction

Background

The intensive care unit (ICU) is a busy working environment where a variety of clinical staff perform different duties at scheduled times of the day, while also having to respond to unexpected, often critical issues with patients. ICUs are typically heavily instrumented, and staff need to be alert to many sources of data from equipment such as ventilators as well as the patients' vital signs and lab test results. For a complex ICU patient (eg, those with multiple conditions), anywhere between 80 and 200 medical interventions are delivered daily and, prior to the COVID-19 pandemic, a member of ICU staff would typically be responsible for up to 10 patients each day [1]. Nonoptimal decisions and clinical errors in this cognitively demanding environment are known to impact patient outcomes [1-4], and a large body of evidence demonstrates that working in an ICU is highly stressful [5,6].

Much of the relevant data for clinical decision-making is already available to staff in the ICU. However, this information is typically scattered across a number of applications, devices, and pieces of paper within the ward. Hence, ICU staff may inadvertently fail to notice signs of a patient's deterioration and struggle to effectively communicate patient updates (eg, test results, medication) or patient requirements (eg, changing tubes, sedative drug management), which will contribute to worse patient outcomes [3,7]. Additional problems such as equipment failures [3] further add to the complexity of working in the ICU and the importance of clear communication among ICU staff [3,8].

The COVID-19 pandemic has generated unprecedented challenges around the globe [9,10], with particularly detrimental impacts on health care systems [10,11]. Increased hospitalizations from COVID-19 put an additional strain on ICU resources, specifically beds with mechanical ventilation [11,12]. In the United Kingdom, this shortage has been such a concern that additional intensive care capacity was made available by the construction of 11 temporary "Nightingale" hospitals [13]. As the pandemic grew in early 2020, the two local ICUs in University Hospitals Bristol and Weston National Health Service Foundation Trust reported a critical need for an information technology solution to help their staff manage increased patient caseloads. The outline brief from the units envisaged a dashboard that would pull together disparate data sources in the ICU to help reduce cognitive load on extremely busy clinical staff. A particular concern was that staff-patient ratios, and hence patient safety, would be eroded by a combination of massively increased patient numbers and COVID-19 cases among their trained staff.

Dashboard for COVID-19 ICUs

The development and use of "dashboards" within health care services are becoming increasingly popular [14]. These are typically interactive, visual tools that help ascertain and monitor trends or the status of key indicators of patients' health condition [15]. We define a dashboard as "a visual display of the most important information needed to achieve one or more objectives [...and] are frequently used to consolidate and arrange these

data so the information can be monitored at a glance" [16]. The key here is that a dashboard must bring together a variety of data and information for (ICU) staff to understand how unwell a patient is and administer appropriate care efficiently and often under time pressure, especially during a pandemic with increasing numbers of extremely sick patients entering ICUs.

It has been consistently found that the use of various graphical displays of information helps detect adverse events and increases clinical diagnostic accuracy [17,18]. Other work has also found that additional clinical support systems bringing together patient information with electronic health records (EHRs) are associated with reduced hospital and ICU mortality rates, shortened hospital stays, and reduced costs of hospitalization (in the US context) [19]. Despite this, it is reasonable to suppose that clinical dashboards could have an important role to play in delivering health care in the ICU, although it is to be expected that this picture will be quite complex. For example, the utility of a dashboard may be limited at moments of crisis since using it requires attention and free hands; new insights into a patient's trajectory might indeed be useful or, alternatively, might introduce additional stress for staff, whereby a specific dashboard may fit the needs of senior staff charged with overseeing the operation of the whole unit, but might be less useful for staff caring for a specific patient [17,20]. However, a recent systematic review and meta-analysis [17] found that dashboards or interactive displays are linked to more accurate, and indeed faster, clinical care decisions in critical care/ICUs. ICU dashboards have been used in Brazil to allocate resources and to obtain near real-time information on suspected and confirmed COVID-19 cases [21]; however, we are not aware of any published study that has interviewed clinical staff and broadly captured the requirements for such a dashboard during the pandemic. Such a study will therefore provide both COVID-19-specific and extreme environment-specific working practices and insights [22]. With new highly transmittable COVID-19 variants emerging and with no guarantee that vaccines will engender an immune response against all future variants, it is hoped that the methodology and the findings in this contribution will be of value to those in a position to develop much-needed new technologies for the clinical frontline.

When building tools and devices, it is important to take a human-centered approach to the design of clinical technologies by including the end-user needs as early as possible [14,23,24]. For reasons of efficiency, patient safety, and job satisfaction, clinical professionals should have a direct role in the design of the tools they will use for their jobs [23-26]. Further, this helps to ensure that critical features in design, particularly in terms of visual displays, are not misunderstood [17]. Therefore, this approach helps to ensure that such tools are developed as a result of clinical "pull," not technological "push," and helps to prevent technology resistance or avoidance [24,27]. Hence, our approach was to commence interviews with clinical staff within the two local ICUs to elicit their requirements for a clinical dashboard. Due to the extreme pressure on staff during the pandemic, the contagious nature of the disease, and the considerable time pressure to implement a solution, these interviews were conducted remotely with only a minimum sample of available

staff with various responsibilities to capture a wide range of requirements.

Since hospitals were dealing with a high demand for ICU beds [28], the aim of this study was to obtain important information to help us understand what design aspects of this tool could help to reduce the workload of clinical staff [4]. Such information could be used to design better tools to help with clinical decision-making processes [4]. Similarly, this may reveal a better understanding of which tasks performed by ICU staff may benefit the most from the introduction of such a dashboard. Further, we anticipate that understanding the various responsibilities of ICU staff could be a useful guide for future participatory design work focused on specific health care staff functions. Therefore, taking into account the preferences of end users for the dashboard design, including how data are displayed (eg, table, figure, graph), enables gaining a better understanding of the role for implementation of these preferences in a multifunctional and customizable dashboard, which could help to prepare ICUs for any future waves and mutations of the virus [29,30]. However, this may require further innovations in clinical processes. We are aware of the complexities of health information systems; therefore, the development and integration of new systems require careful consideration and human-centered design [23]. If this is not the case, there is a risk of developing new systems surplus to requirements, or causing further complications if staff need to switch between multiple systems. New systems could potentially compromise patient safety, be difficult to use and learn, and encounter resistance from staff, potentially resulting in poor uptake of the system [23,26].

Hence, the overall aim of this project was to collate a series of insights and requirements from end users across two ICUs to

design and build a clinical dashboard to support their increased workload during the COVID-19 pandemic. Requirements elicitation, as seen in this work, by its nature is difficult, as requirements are volatile and necessitate translation from natural language via stakeholders through to tangible software [31,32], especially with a fairly unknown disease that continues to change, mutate, and impact society. The current pandemic has been an “extreme environment” [22] for researchers across disciplines. Hence, this work provides interesting insights regarding COVID-19 specifically but also provides an example of work conducted “in the wild” that encompasses the context, nuances, and uncertainties faced by both the researchers and the ICU staff [33]. The difficulties range from the time and additional pressure ICU staff were under at the time of writing but also entire workforces required to “work from home,” thus diminishing the inability to travel and collaboratively work together to previous expectations. Altogether, this provides a novel frame for this research to occur from initial planning, interviews, through to subsequent testing evaluation, and future iterations being deployed as this work is ongoing.

Methods

Research Questions

At a time of unprecedented workloads in the ICU, clinical staff time was in short supply. We interviewed six staff members working in ICU wards across two hospitals for the UK National Health Service (NHS) (Table 1). We were acutely aware of the additional strain on the NHS and staff; hence, we proceeded with interviews as a direct and simple method to capture requirements and reduce additional workload on ICU staff [31,32].

Table 1. Participant job role across the two intensive care units.

Participant ID number	Job title	Hospital
1099	Consultant in Intensive Care Medicine	A
1159	Sister (Band 7, Manager)	B
1252	Anesthetic Registrar	B
1587	Consultant in Intensive Care Medicine	B
1704	Consultant in Intensive Care Medicine	B
1839	Matron (Senior Nurse/Nurse Manager) in Intensive Care	A

Our exploratory interview questions were:

1. What tasks do each participant perform in the ICU, and where would a portable dashboard be helpful to improve safety or reduce workload on staff?
2. Do COVID-19 patients pose particular clinical challenges that need to be taken into account in the dashboard design?
3. Which data would staff need to look at on the dashboard and how should they be presented?
4. Would staff have any concerns about the introduction of a portable dashboard into the ICU?

Participating Hospitals

The two hospitals participating in the study were distinct from one another, with one (hospital A) having a more technologically enhanced ICU where many systems are already digitized in comparison to hospital B, which remains more paper-based with a much larger ICU unit. The interviews were semistructured, lasting approximately 30–45 minutes, and due to the pandemic situation in early 2020 were conducted exclusively online. Topic guides were used to assist the interviewer and encourage consistency, with additional post hoc questions to further explore any potential themes. The interviews were recorded and transcribed verbatim. We used

NVivo (version 12) [34] for the qualitative coding process and conducted an inductive thematic analysis [35].

Analysis

To analyze these interviews fully, we took a multireasoning approach within our thematic analysis, which consisted of both deductive and inductive approaches to ensure rigorous development of coding trees [36,37]. Initially, we developed a preliminary codebook deductively via the data familiarization phase. We then entered a second phase that took an inductive approach to allow new themes to materialize as we coded each interview. We iterated through this process until the researchers reached a unanimous decision on the final codes, themes, and how they fit together. This in-depth and rigorous methodology aimed to ensure we were as thorough as possible for capturing requirements, alongside taking advantage of our multidisciplinary team, noting that the qualitative coders were not medically trained unlike other members of the team. Hence, it was important for us to ensure that we were accurate and appropriate in our codes, themes, and understanding [38,39]. Therefore, we sought medical insight and advice at every stage of the process within our team. We produced two distinct codebooks: one relating specifically to the requirements of a dashboard and another that related specifically to concerns of dashboard use in ICUs. The requirements codebook consisted of 96 codes that were initially subdivided into the following three requirement categories: 24 technical codes, 56 clinical codes, and 16 operational/logistical codes (see [Multimedia Appendix 1](#)). The concerns relating to the dashboard use codebook consisted of 24 codes, including 9 codes regarding design and 15 codes concerning operations (see [Multimedia Appendix 2](#)). The final codebooks were then used as the guide to code the qualitative interviews; therefore, in NVivo, each code would have words, phrases, and quotes from participants organized into these codes and themes. This ensured a straightforward and well-organized analysis. Hence, our hybrid approach [37,40] was the method employed for this codification, which identified existing patterns and subsequently regrouping codes into their emerging themes.

Ethics Statement

This work was approved by the Faculty of Engineering Research Ethics Committee at University of Bristol (case 2020-3236).

Results

Overview of Key Requirements

Based on our interviews, we elicited five key requirements from a range of ICU staff to capture a wide range of roles and needs when using clinical dashboards. For instance, a dashboard must be adaptive and flexible to continue to be useful in changing clinical environments. Furthermore, dashboards need to be customizable because different staff may have specific parameters and information they need to see first (eg, “condition at a glance”). This can be achieved by ensuring there is some customizability for individuals or staff groups to select the information that is displayed. A dashboard would need to be mobile, which could reduce reliance on paper-based forms with patients. This also relates to task and patient management, where

several staff noted that a dashboard could help with data entry and management alongside assisting with patient handovers, as all information will be collated into one system that is easily accessible. In contrast, some concerns were raised, for instance around infection control when carrying devices in and out of ICUs and high-risk-of-infection areas.

Requirements for an ICU Dashboard

Flexibility With Changing Protocols for an Evolving Disease

Unsurprisingly, the new challenges posed by the COVID-19 pandemic featured heavily in the participants’ responses to our interviews (see [Multimedia Appendices 1 and 2](#)). As mentioned above, as these interviews were conducted in early 2020, relatively little was known about the nature of this disease and how to contain or treat it. This had direct implications for hospitals, highlighted by Interviewee 1587 (ICU Consultant), who described ICU staff difficulties handling COVID-19 patients while “keeping on top of constantly changing best practices.” At the time, no consensus had been reached regarding the clinical management protocols for COVID-19 patients, with staff under pressure to run additional tests while keeping track of the results. According to Interviewee 1704 (ICU Consultant), this absence of clear protocols created a very difficult working environment for staff:

...for a couple of days when I was on the COVID-19 side, full witness to it, it was a bit of a shambles and very stressful for the staff. There was no harm done to any of the patients, but it was asking staff to work outside their comfort zone and people found that professionally very difficult...

As ICUs are data-intensive environments by their nature, while in the process of adapting to a new disease and ever-changing protocols, a dashboard could help avoid staff forgetting to check certain parameters or simply avoid them due to being overwhelmed with new tasks and information. For instance, Interviewee 1099 (ICU Consultant) stated: “[a dashboard has the ability to] draw[s] my attention to things that I would probably forget about if I’m honest. I can only process X amount of information.”

One of the great challenges for medics and hospitals during the pandemic has been keeping on top of the new, fast-moving information and updates regarding COVID-19 management and treatment:

We are still learning a lot about [COVID-19 patients]. When [information] was coming from Italy, there was a lot of talk about how the patients were and how we were supposed to treat them, pretty much everything we were told has been wrong. It seems to be a very unusual disease and it’s not like anything we have seen. Whilst we were told it was a really bad pneumonia, which we had to treat with aggressive ventilation; it turns out it seems to be a disease of blood clotting, that affects the lungs. It makes our treatment that we are doing the wrong thing and potentially even harmful, so we’ve changed a lot

about what we are doing... [Interviewee 1704, ICU Consultant]

It's a new disease process. We are learning all the time; best practice of evidence is constantly changing so keeping on top of that is difficult. [...] Accessing best practice can be difficult, we have a new intranet [...] where] we have our single page checklist with guidance for these patients that is shared with [redacted]. I was thinking about what sort of things would be helpful and actually because it is a new disease process there is quite a lot of new things we don't normally do, so a lot of regular blood tests that happen for example at days 1, days 3, 5, 7, keeping on top of when those are. They are all on our daily checklist, on our management guide. They are tests we wouldn't normally do, but they are looking for specific things. Other new disease processes on top of the COVID, like HLH [hemophagocytic lymphohistiocytosis] that can happen, so sort of screening for those. [Interviewee 1587; ICU Consultant]

Interviewee 1704 (ICU Consultant) further noted the continuous operational changes occurring in the COVID-19 “pods,” defined by Interviewee 1587 (ICU Consultant) as the “designated COVID-19 areas” where additional personal protective equipment (PPE) is required for entry, which created additional psychological and physical strain:

The COVID pod has changed quite a bit because advice on personal protection has changed. Initially, all the patients were in closed rooms with the doors shut and you had to put all the equipment on to go in and see them, [...]. Now, everybody is wearing it all the time in the hall throughout the unit and the doors are open. So, when the doors were shut you didn't do the ward round at the bedside, you were stood outside looking at all the data and the nurse inside was writing stuff on a white board to show somebody on the outside to write on the piece of paper. That wasn't a sustainable solution because it was quite staff intensive.

[when health care staff move into the] high-risk areas [...] there is a barrier psychologically to going into the room.

The accounts of participants illustrate the importance for an ICU dashboard to have the capacity to be flexible and easily updated due to constantly changing protocols, information, and advice (eg, reminders to check for specific parameters, changes in protocols regarding closed rooms or designated COVID-19 pods). This could draw from national- and hospital-level advice or information that is pushed out to staff via the dashboards pulling from various NHS information systems. This is a key requirement to ensure the dashboard continues to be usable as we learn more about COVID-19 and to potentially adapt this device into a tool that can remain integrated into ICUs more generally. These narratives highlight the ICU staff needs for an adaptable dashboard that can be updated with constantly changing real-time data about patient parameters, with new and revised routine alerts for new tests, and with reminders for

specific trends to look out for when dealing with COVID-19 patients.

A Mobile Dashboard

Interviewees 1252 (Anesthetic Registrar), 1587 (ICU Consultant), 1704 (ICU Consultant), and 1839 (Matron) stated a preference for a dashboard that they would be able to use while walking around the ICU to attend to patients:

There is a lot of walking around by the nurse in charge, just to touch base with people for support and things. They are not static, so it would have to be a mobile solution [Interviewee 1704, ICU Consultant]

A lot of my clinical duties are mobile so to not have the technology follow me and having to use fixed desktops is sometimes quite frustrating. As I say I am mobile, the whole of clinical care is conducted in a very mobile fashion [Interviewee 1252, Anesthetic Registrar]

Thus, the dashboard would potentially replace paper forms and allow for simple and efficient data entry (both qualitative and quantitative) as staff move around while on shift. Interviewee 1704 (ICU Consultant) discussed new issues that arose specifically from dealing with COVID-19, such as the heavy reliance on paper forms in their hospital, given that computers are situated in high-risk areas, which, due to PPE use regulations, makes access to computers difficult. Hence, having a dashboard with remote access to various hospital systems and records that can also be used in mobile devices and taken by staff members outside of the pods would be helpful.

The other thing that COVID-19 has bought in with it, is our clerical staff [...] who input a lot of data onto the computer, they're not able to go into the area because they're not fit tested with the masks. We are taking pictures of the observation chart with the iPads and uploading them to [redacted] [...] for them to then look at remotely, which is a bit of a fudge. It's not brilliant, but it's better than nothing at the moment.

By reducing paper forms and collating information neatly in one place, this would ensure information can be carried from patient to patient around the ward and would simplify a series of traditionally offline protocols, task, and data management. While simple in concept, realistically bringing together several diverse information systems and data across hospitals into a unified system or database is a highly complex task and may be difficult to integrate fully into the workforce [41]. However, examples of publicly available, deidentified EHRs do exist, such as the Medical Information Mart for Intensive Care (MIMIC) as part of the Beth Israel Deaconess Medical Center [42].

Customizability and Usability

A key theme arising from the interviews was the users' needs to customize the dashboard (see [Multimedia Appendix 1](#)). For example, Interviewee 1099 (ICU Consultant) expressed a need for customization to support the different tasks and roles of the ICU staff, since clinical information such as patient parameters

is crucial to performing the Intensive Care Consultants' tasks, whereas operational data such as duration of patients in prone position would not necessarily be of interest to doctors yet would be of great importance for nurses. Similarly, Interviewee 1252 (Anesthetic Registrar) suggested the usefulness of knowing the patients' pending and past procedures. Hence, the data required for a dashboard to be useful are extensive, complex, and would draw from several hospital information systems, including (where * denotes parameters that were also mentioned as important to view over time as trends and ^ denotes markers particularly important and of interest for COVID-19 patients): C-reactive protein (inflammation marker)*^, D-dimers (blood clotting marker)*, ferritin (inflammation marker)*^, lymphocyte count (inflammation)*^, platelet count (inflammation)*^, procalcitonin levels (inflammation)*^, pending tests for specific HLH patients, blood pressure*, white blood cell count*, fraction of inspired oxygen, oxygen level*, oxygen supply level for personalized care, peak airway pressure for COVID-19, tidal volume size, positive end-expiratory pressure^, arterial oxygen/inspired oxygen ratio (PF ratio), plateau pressure^, type of ventilation, Sequential Organ Failure Assessment (SOFA) score, Glasgow Coma Scale score, intracranial pressure*, ventricular tachycardia^, infusion rate of vasopressors*, number and absorption of nutrition calories, and COVID-19 status. Further, specific requests for alerts to be associated with specific data were noted, such as abnormal values across parameters, 7-10 days of static oxygen (meaning a computed tomography pulmonary angiogram scan can be completed), pending tests for COVID-19 patients, clinical deterioration, and pending procedures for patients. Finally, other information was requested that was more operational in nature, including bed layouts, number of patients on dialysis, nurse locations, patient numbers across units, patient flow information (eg, admission, discharge, changing of units), and relevant patient handover information.

Unsurprisingly, a lot of data and information were requested to appear in the dashboard, which may be overwhelming and difficult to navigate or simply not of relevance or interest to certain roles in the ICU. Interviewee 1099 (ICU Consultant) suggested that a fully individualized dashboard for each staff member would be ideal as this allows for a tailored configuration for each of the roles within the ICU. However, it is important to be cautious when implementing highly customizable systems, since this customizability may induce errors when overstressed staff are required to make fast decisions. Hence, careful consideration is needed to determine what degree of customization is advisable for this device at the user level. It is important to have functionality that allows additional staff access on an ad hoc basis. As Interviewee 1587 (ICU Consultant) stated, staff from different areas of the hospital may need access to the dashboard. For example, COVID-19 patients often need nutritional and dietary assistance; hence, nutritionists may also need specific systems access for these patients.

Alternatively, Interviewee 1839 (Matron) suggested having a split view where parameters could be presented as broader categories such as "Clinical Parameters" (eg, ventilation, tidal volumes, dialysis) or "Safety Parameters" (eg, delirium, infection, prone position turns), so that these measures could be useful for various health care staff functions and

responsibilities. However, this could also be problematic if functionality allows for toggling role-specific parameters and the lack of prima facie data may cause staff to miss trends in parameters not shown on the screen.

Several participants noted the importance of integrated systems (see [Multimedia Appendix 1](#)). For example, Interviewee 1099 (ICU Consultant) stated their current systems necessitates that they "have to open up 5 different screens to get the data [they] need and that is pretty labor-intensive," or that they are "often chopping and changing through different programs..." which demonstrates the importance of having an intuitive interface that consolidates relevant data on demand. Similarly, in terms of usability, it is problematic if a system requires additional "administration" for staff to find information and results on pending tests. This can be illustrated with the following account from Interviewee 1099 (ICU Consultant) about the [redacted] system:

When you have somebody who comes in with a chest infection, 15 tests are ordered: five aren't back, five were never sent, and five are back, but you would never know that by looking at [a system]. You would only know that by looking at this separate program.

It is therefore crucial for the dashboard to be well-integrated with other hospital systems to avoid data and work duplication. It is thus paramount to ensure that a new dashboard does not add complexity but reduces workload to access data by extracting it from the existing systems. It is essential for staff to be able to customize their view to quickly sift through large amounts of information, understand patient needs, and determine next steps. As a requirement, this is achievable and realistic, as it is common for information systems to have user profiles with individual logins [43]. Further, having default user profiles based on roles and grades within the workforce is a reasonable requirement to implement, where individuals can request additional accesses ad hoc. However, the level of customizability offered alongside user profiles would require additional testing and research (eg, changing color schemes, data access changes).

Dashboard Layout and Trends of Incoming Data

Patient Overview

In general, there was a tendency for all participants to comment on how data should be processed and presented on the dashboard. Interviewee 1839 (Matron) stated a preference for the ability to view the whole ward (which could also track bed and patient expansion). If this were to match the physical layout of the beds, this would be useful to find patients quickly and to effectively plan a patient's acuity (level of nursing care) quickly. This is especially important should the unit become busy or indeed require a quick and large-scale expansion in patient numbers, as envisaged at the UK's "Nightingale" hospitals. In addition to the ward view, Interviewees 1099 (ICU Consultant), 1704 (ICU Consultant), 1839 (Matron), and 1252 (Anesthetic Registrar) suggested the inclusion of a summarized "Condition at a Glance" view, which would allow for those starting their shift to quickly get up to date. This was echoed by Interviewees 1704 (ICU Consultant) and 1839 (Matron), who suggested a display of "overarching" parameters of all patients on the ward.

Interviewee 1587 (ICU Consultant) reported a similar requirement that would display the most critical parameters (eg, SOFA scores, tidal volume), which would allow the Intensive Care Consultants to see the trend of a patient's current condition. Interviewee 1587 (ICU Consultant) argued that calculating SOFA scores is an arduous task for junior doctors; hence, making these calculations available and easy to interpret in a dashboard will save ICU staff time and cognitive energy [44].

Of course, there is more to ICU patients' health care than a first-glance interpretation of these parameters. Interviewee 1839 (Matron) stated the complexity of patient monitoring when patients might appear fine in terms of typical baseline metrics, but in reality, their actual state is misrepresented by data: "[a] patient could be fine, but they're on a lot of inotropes (eg, noradrenaline) or ventilation, and it looks OK, but they are on 100% oxygen and/or quadruple noradrenaline." The intuitive assessment required in such cases could only be achieved through an appropriate identification of the daily and hourly trends by the ICU staff, of the patients' inflammatory markers and oxygen levels, according to Interviewees 1704 (ICU Consultant) and 1587 (ICU Consultant). Thus, the availability of these data on a dashboard will be particularly important for health care staff to decide on COVID-19 patients' clinical care and to devise provisions for their safety, as expressed by Interviewees 1099 (ICU Consultant) and 1704 (ICU Consultant). Further, Interviewee 1587 (ICU Consultant) discussed the significance of data trends, especially regarding COVID-19 patients, such as decisions about when patients can be weaned off ventilation as well as monitoring the prone position ventilation timings. Additionally, a dashboard could track specific sets of data, which may offer statistical insights (Interviewee 1252, Anesthetic Registrar) into how to better treat future patients for particular diseases or conditions, as these data will be able to provide baselines and expectations, particularly for new diseases such as COVID-19.

Data Visualization, Warnings, and Alerts

There were differences among ICU staff as to how they wanted data to be presented in this type of tool (see [Multimedia Appendix 1](#)). A suitable example of this would be the informants' preferences for graphical displays that would help to address two important issues pointed out by Interviewee 1587 (ICU Consultant):

I think if you click on it, you can see a graph, but to be honest we don't regularly do that. Ninety percent of the time it's just numbers completed on a sheet... We literally have a piece of paper that junior doctors fill in in the morning or if they haven't, we go on [redacted] and click on the CRP trend on there and see what's happening. You can get a graph of it, but it's not ideal.

Hence, it is important to enable users to have a degree of autonomy to customize the dashboard for individual patients (Interviewee 1099, ICU Consultant), and to transform data and information into a format that best suits their learning needs and information processing style. Alongside ways to visualize data, it is critical to have suitable real-time deterioration alerts both for clinical and safety parameters, which may include

visibly highlighted alerts on abnormal values, as well as real-time alerts on staff deviations from practice (Interviewees 1099 [ICU Consultant] and 1704 [ICU Consultant]). However, when it comes to the display of these warnings, Interviewee 1099 (ICU Consultant) discussed what would be the most suitable parameters for COVID-19 patients (eg, PF ratio and driving pressure) to help clinicians decide when to start weaning the ventilators, as well as the adequate time parameters for the graphic view. This interviewee stated the importance of including suitable "cutoffs" in terms of data presentation, because to plan COVID-19 patient care, Intensive Care Consultants must consider both the presence of abnormal values as well as how these values behave over time (trends). In addition, Interviewee 1839 (Matron) noted determining thresholds for colors and notifications is not a trivial matter, especially if they need to be tailored to specific medical conditions or personalized to each individual patient. Therefore, a dashboard that presents data in a way so that the user could see a longer patient history may be extremely helpful. This could be facilitated by allowing the graphs to be scrolled through horizontally to show earlier data and see longer longitudinal trends since the patient's admission.

It is important to be aware that notifications or alarms are extremely common in ICUs due to the variety of abnormal values of health parameters in critically ill patients. Therefore, when designing a new dashboard, a reasonable balance is needed to avoid "alarm fatigue" and to prevent staff from missing patient deterioration markers, which can lead to detrimental outcomes [45-48].

The dashboard would give me the triggers to go sniffing around into the detail of the data [...], I would just highlight the noradrenaline box and the base excess [and hide other parameters and] look at those two things. What's the trend? [...] I think the personalization of being able to manipulate it on one screen... [Interviewee 1099; ICU Consultant]

trends are brilliant. [...] it doesn't really matter what the noradrenaline is, if it has doubled in the last hour, it's not a good thing. [...] Actually, it's the step change that is the important thing. For me, I was straight drawn into the color change and the arrows, then just lost sight of the numbers a bit. That is probably a good thing because I would then go looking into that patient detail on the system to see why their noradrenaline is going up and doubled. So that was quite a good trigger. [Interviewee 1099; ICU Consultant]

Here, the requirement relates to the layout of the dashboard and how data are presented (with some in real time). This called for flexibility in terms of how data can be graphically displayed to suit the staff member using the device, which may also help with the number of alarms and notifications in the ICU. This requirement is relatively simple to implement within a dashboard system, where users will be able to shift between graphic and tabular displays of information or seeing longer-term trends of a patient, for example. However, attempting to visualize data and highlight when parameters are shifting negatively for the patient is inherently more complex; thus, testing and examining

what the thresholds should be are crucial to reduce alarm fatigue among additional stress for staff. This would include investigations regarding individualistic measures versus overall baseline “cutoffs.”

Task and Staff Management

Patient Handovers

A key theme regarding staff and task management were patient handovers (see [Multimedia Appendix 1](#)). Handovers (both regarding staff shifts and turnover of their allocated patients), are complex two-way processes between a variety of staff signing out and updating those coming in to take over, where the accuracy and effectiveness of this information exchange “will facilitate consistency and continuity of care” [49]. This is particularly important for critical care patients, where omission of pivotal information during the handovers could influence future treatment and subsequently cause failures in patient management [49]. Since the use of a dashboard could rapidly help to capture and track wider information regarding patient status and care requirements, it is evident that the implementation of this tool in the ICU environment could facilitate more structured and effective patient handovers. For example, Interviewee 1839 (Matron) stated:

We know that handover time and transfer of care is a pinch point where if there is going to be an issue or problem occurring, we often track it back to that point in time. Where something has been missed, not handed over or at that point they may look at something and go “that’s not what I remember it being.” That’s the trigger to go back systematically through all their different bits. Or a doctor has come along and changed the rate of a pump and not told somebody. We know that is a really pivotal time so some kind of overarching view of the main clinical elements of a patient care would be helpful. That would give them a visual aid to that and anything that would help a hand over of care, would absolutely be welcome.

Similarly, Interviewees 1099 (ICU Consultant) and 1839 (Matron) coincided in stating the importance of allowing time for scrutiny of new inputs of their colleagues and data updates from various patients that have been handed over. However, as Interviewee 1099 (ICU Consultant) pointed out, outlining changes in patient parameters during quick handovers is done with great difficulty while having to navigate multiple hospital systems to gather the information required:

I had a handover from my colleague - but I want to process it in my own mind and want to see what’s changed over the last 12 hours, since they handed over, [currently] I have to open up 5 different screens to get the data I need and that is pretty labor intensive.

Data Entry

Due to the important consequences at stake, extreme care and monitoring are exercised in the ICU environment to ensure accurate data input in their systems.

If you are feeling responsible for the patient, which you are as a consultant, you need to double check that [data, patient notes]. The only way to do that is to physically look at those [systems] and paper notes yourself [Interviewee 1099, ICU Consultant]

Moreover, when planning patient care targets, ICU staff in managerial positions must carefully balance patient-management workloads of staff with their data input tasks. This links in with comments from Interviewee 1099 (ICU Consultant) when expressing concern about overtired staff with data entry responsibilities such as recording general observations about patient progress and invasive procedures, among others:

[staff] will be absolutely knackered at three in the morning and just put [in] the bare minimum. They’ll go, “patient had an operation, and this is what happened” and forget about other stuff.

On these grounds, participants considered having a tool that can assist staff in the transcription and modification of patient data with a minimum error rate to be important. Arguably, errors may continue to occur in the presence of a digital interactive dashboard; however, research has shown that using digital systems to collect and log data (rather than pen and paper) reduces errors in data recording and data entry [50,51]. The use of a dashboard helps to provide a faster way to populate handover or debrief notes. Furthermore, Interviewee 1252 (Anesthetic Registrar) noted the lack of any formal system at present to document any medical advice provided to patients over the phone. This could be a simple note about the patient that can be added to the dashboard to ensure an overview of all advice and information previously provided to the patient.

Task Management

A major issue reported by the participants was a lack of warnings in their current systems about forthcoming completion times of pending tasks and targets, which could be built into a dashboard. Interviewee 1099 (ICU Consultant) stated that a careful balance must be struck with off-target warnings to avoid undermining staff confidence:

The warnings that are built into the target need to be in advance, there is no point telling people at midnight you’ve not met your fluid balance target because it will just demoralize people.

Consequently, it is unsurprising that one of the participants’ most frequently mentioned dashboard requirements was to have warning notifications ahead of completion times (see [Multimedia Appendices 1 and 2](#)), which would certainly work as a task management system that will ensure timely completion of the multiple pending tasks and daily targets of medical and nursing ICU staff, including ventilation weaning, daily prone and supine ventilation sessions for COVID-19 patients, and other safety tasks of nursing staff; invasive procedures such as tracheostomies; monitoring pending microbiology and specific tests for COVID-19 patients; changing of drugs; and speaking to relatives (Interviewees 1099 [ICU Consultant], 1159 [Sister], 1252 [Anesthetic Registrar], 1587 [ICU Consultant], 1704 [ICU Consultant]), among many other responsibilities. Interviewee 1099 (ICU Consultant) stated that this type of task management system would give managerial staff peace of mind by knowing

that “loops are closed,” especially when the ICU becomes extremely busy and “people forget about minutiae.” Furthermore, in a context of ever-changing guidance and information, Interviewees 1159 (Sister) and 1587 (ICU Consultant) also agreed that it would be extremely useful to have daily task checklists (eg, safety checks) as a suitable requirement of the dashboard, which should also include enabling inputs of data as needed. Interviewee 1587 (ICU Consultant) also pointed out how having a dashboard to prompt staff members to finish tasks would be important to address the absence of warnings on pending targets in their current system, especially with new diseases such as COVID-19 when more tests than usual are frequently needed. For example, Interviewee 1099’s account illustrates how ICU staff struggle to juggle their immediate tasks with their daily patient care targets for both COVID-19 and non-COVID-19 patients:

There will be a patient that will be on multiorgan support, have to go for a scan, have to go to theatre, come back. Then the nurses will try to make sure the patients are fed, [...], all sorts of complex care issues [are] going on. That [eg, fluid] target then drifts into the background. There’s no prompt to say, “your fluid balance is nowhere near target and you have four hours to go. What are we going to do to solve this problem?”

To obtain a general picture of the ICU patient care flow, interviewees discussed the advantages of having an overall view of staff numbers and their corresponding workload, alongside data about patient admissions and patient flow (eg, patient discharge and transfers either for tests or to other wards), as pointed out by Interviewees 1159 (Sister), 1839 (Matron), 1704 (ICU Consultant), and 1252 (Anesthetic Registrar). For example, the following testimony of Interviewee 1159 clearly portrays the managerial ICU staff need of a well-updated and integrated system showing workload allocation:

We need a way of knowing which nurse is in each bed space so if there is an issue, we can speak to that nurse looking after that patient. If it was two shifts down the line, there was something we needed to get hold of somebody about.

At the same time, from the following account of Interviewee 1704 (ICU Consultant), it can be inferred that managerial staff are simultaneously responsible for overseeing the staffing of the unit while monitoring the changing conditions of all patients. Hence, the requirement for a new system to collect these two categories of data was mentioned:

They collect this data on a piece of paper as well and it’s in pencil so they can rub stuff out and change it. I have been trying for years to get iPads for them to use, we need some kind of software for that. That data would be really important to analyze: the patterns of activities during the day, to optimize our staffing models or things like that.

This statement highlights that the implementation of a dashboard for ICU use could provide a much-needed opportunity to shift from pen and paper to a digital system of data collection and monitoring, alongside a new strand of data analysis that could

help optimize staff time and workload. Interviewee 1252 (Anesthetic Registrar) provided additional corroboration of a dashboard’s value to mobile staff such as Registrars by having an effective oversight of all patients pending transfers to other wards and for tests (eg, for a computed tomography scan or other tests). This relates to Interviewee 1704’s (ICU Consultant) statement about the need to better optimize staffing models and the daily distribution of tasks by accurately monitoring staff workload and whereabouts as follows: tracking staff timings for patient care for each allocated patient, as well as producing continuous insight into the location of the health care staff throughout the different ICU wards during their shifts (Interviewees 1159 [Sister] and 1252 [Anesthetic Registrar]). According to Interviewee 1252 (Anesthetic Registrar), this two-fold patient-staff tracking system would be very useful for staff who are constantly busy (Registrars and Running Nurses). With this new system, they can efficiently share their work in relation to their location, the patient transfer destinations, and the numbers of daily transfers.

This final requirement, which touches on all prior requirements for the dashboard, is ensuring that the dashboard is seamlessly integrated with other hospital systems. In this way, staff can access additional, external data (eg, authorized views of blood test results from other hospital systems). This would help to tackle the issue described by Interviewee 1252 (Anesthetic Registrar) when characterizing the process of patient admissions from other hospitals as “data heavy” with a lot of “transcribing of various different sources onto the intensive care unit systems,” where data are currently not being pulled neatly into one system. As stated previously, this is a highly complex task that would be difficult to integrate as current systems are siloed across hospitals [42].

Concerns About Dashboard Use

As shown in [Multimedia Appendix 2](#), there were two main operational concerns among ICU staff regarding the use of dashboards. The first concern, raised by Interviewee 1587 (ICU Consultant), related to the potential increase of infection via use of mobile technology and equipment. On the assumption that these dashboards will be used in mobile devices such as iPads, Interviewee 1587 (ICU Consultant) questioned whether these devices should be allowed into the infection-controlled areas of the ICU, specifically at the bedside of COVID-19 patients. As possible solutions, Interviewee 1587 (ICU Consultant) mentioned the use of disinfecting wipes for mobile devices, but mostly adopting clinical protocols that avoid the need for using a device at the bedside.

We don’t take the iPads in to see patients. In terms of risks, they are potential fomites, a sort of vector for transmission of infection. We wouldn’t take them into bed spaces. Much like cleaning our phones, we are good at cleaning with special Clinell wipes. Yes, there is a potential risk... We stand outside where it is lower risk. We cluster round as a ward round to write up our notes and decide the plan. Only one of us will go in to examine the patient. Everyone else will wait outside.

The second operational concern regarding dashboard use related to the differing levels of technology literacy among ICU staff. For example, Interviewee 1099 (ICU Consultant) stated that unless there already is a culture around the use of dashboards and technology, encouraging staff to actually use and engage with this type of tool might be difficult [27]. This is an interesting comment as Interviewee 1099 is based at the more technology-enhanced ICU that uses dashboards among other devices regularly, where such tools are integrated. Hence, these concerns additionally relate to the managing expectations in staff (Interviewee 1587, ICU Consultant) of what the dashboard will do, how it should be used, and protocols regarding these devices, as it is important not to oversell a new technology's potential impact for the ICU.

Finally, there was some ambivalence from staff members about the dashboard design, as illustrated by the concerns raised by Interviewee 1099 (ICU Consultant), who is based at an already technology-enhanced ICU, in relation to what should be the "acceptable" parameters for the notification and warning timings with regard to both COVID-19 patients and staff targets. Interviewee 1099 (ICU Consultant) also raised concerns regarding notifications or alerts coming in at inappropriate times, since this could cause the adverse effect of "demoralizing people." This participant further mentioned a concern for having a tool with which staff could compare their unit target achievements with other units. This could increase ICU staff stress, and might lead to suffering from burnout or cognitive overload, "wow, mine's [targets or parameters] all red, things aren't going very well," which could have ramifications for both patients and staff due to the well-documented fact that the ICUs are incredibly stressful environments [5,6,52]. These are important concerns to address early and to ensure staff are all fully informed regarding the system itself and the transparency regarding how the digital logs that it will produce may be used outside of patient monitoring (eg, can these be used to assess staff performance in the workplace?). Hence, when implementing new systems, the engagement of end users is key to ensure expectations are set and staff can feel supported by these new systems.

Discussion

Principal Findings and Conclusions

In response to the critical situation of two local ICUs, we conducted a series of interviews to elicit requirements for a bespoke dashboard to help ICU staff save time and work more efficiently, particularly during the COVID-19 pandemic. We found that despite having limited access to end users, our approach of conducting remote requirements interviews for developing a dashboard for COVID-19 ICUs has been successful. The rapid cycle of interviewing end users, prototyping the user interface, and iterating over the software design, despite taking place in extreme and distressing circumstances of the pandemic, has proven to be an effective way of producing functional software requirements. These requirements have in turn allowed for the development and deployment of an interactive dashboard currently being tested and evaluated across two hospitals.

The first requirement was the need for a flexible dashboard, primarily to help ICU staff respond to rapidly changing guidance for the management of this new disease. The second requirement emphasized the need for a mobile dashboard, which allows staff to walk around wards with real-time data and information of patients. The third requirement focused on customizability of a dashboard, stemming from the great diversity of roles and tasks conducted by ICU staff. Related to this was the fourth requirement, which was the ability to track and visualize real-time data and daily/hourly trends on patient parameters. The fifth requirement was aimed at pending tasks and targets for staff management. All requirements highlight a need for the integration of different hospital systems within the dashboard, which is a longstanding challenge in medicine [41]. Alongside these requirements, participants raised concerns regarding the infection-risk safety issue of bringing devices into the ICU and of the timing of warnings and alerts.

The study findings confirmed that digital solutions for ICU use would potentially reduce the cognitive load of ICU staff and reduce clinical errors at a time of notably high demand of intensive, critical health care [17]. As summarized by Interviewee 1099 (ICU Consultant), the beneficial implications of having this dashboard would hopefully be that "not only will it make the system more efficient" but it will further give them the possibility of "looking after more patients more safely."

Limitations

We acknowledge that the sample size is small due to the workload on ICU staff caused by the COVID-19 pandemic, which was the underlying motivation for this research. However, we did capture requirements, perspectives, and experiences from a wide range of clinical roles within the ICU environment across two somewhat different hospitals, particularly from those heavily involved in the health care of COVID-19 patients during the pandemic. There are a number of ways we could have elicited the requirements for the dashboard, including from questionnaires, joint application development, storyboarding, and protocol analysis [31,32]. However, with the time pressure to develop a dashboard that was working and usable, alongside the time pressures ICU staff were under, we wanted to continue with the most straightforward and least cognitively heavy method of elicitation for the ICU staff.

Further, we are aware that our dashboard will require extensive testing in the ICUs with end users such as our interviewees to refine the design and functionality. This would include examining how we would tailor the dashboard for different roles (eg, having a home page with various types of information for each role such as nursing staff vs a consultant or a nutritionist). This will be an iterative process, where we acknowledge that not all needs will necessarily be met; however, our aim is to ensure the device is usable and enhances staff.

We are also aware that our sample comes exclusively from the region of Bristol, which may not be representative for ICUs across the rest of the United Kingdom or indeed outside of the United Kingdom. However COVID-19 has impacted health care provision in many regions and many countries worldwide, and many of the staff, patient, and task management requirements; the ability to track and monitor trends; and the

dashboard customization for individual staff members are likely and around the world [53].
to be common requirements both across the United Kingdom

Acknowledgments

This work is partly supported by the Engineering and Physical Sciences Research Council Digital Health and Care Centre for Doctoral Training at the University of Bristol (UKRI grant EP/S023704/1 to MW). CMW is funded by the South West Better Care Partnership, supported by Health Data Research UK.

Conflicts of Interest

None declared.

Multimedia Appendix 1

Code tree classification of dashboard information preferences (96 codes).

[\[XLSX File \(Microsoft Excel File\), 12 KB - humanfactors_v9i2e30523_app1.xlsx \]](#)

Multimedia Appendix 2

Code tree classification of concerns about dashboard use (24 codes).

[\[XLSX File \(Microsoft Excel File\), 10 KB - humanfactors_v9i2e30523_app2.xlsx \]](#)

References

1. Pronovost PJ. Enhancing physicians' use of clinical guidelines. *JAMA* 2013 Dec 18;310(23):2501-2502. [doi: [10.1001/jama.2013.281334](https://doi.org/10.1001/jama.2013.281334)] [Medline: [24310916](https://pubmed.ncbi.nlm.nih.gov/24310916/)]
2. Cullen DJ, Sweitzer BJ, Bates DW, Burdick E, Edmondson A, Leape LL. Preventable adverse drug events in hospitalized patients: a comparative study of intensive care and general care units. *Crit Care Med* 1997 Aug;25(8):1289-1297. [doi: [10.1097/00003246-199708000-00014](https://doi.org/10.1097/00003246-199708000-00014)] [Medline: [9267940](https://pubmed.ncbi.nlm.nih.gov/9267940/)]
3. Reader TW, Flin R, Cuthbertson BH. Communication skills and error in the intensive care unit. *Curr Opin Crit Care* 2007 Dec;13(6):732-736. [doi: [10.1097/MCC.0b013e3282f1bb0e](https://doi.org/10.1097/MCC.0b013e3282f1bb0e)] [Medline: [17975399](https://pubmed.ncbi.nlm.nih.gov/17975399/)]
4. Norman GR, Monteiro SD, Sherbino J, Ilgen JS, Schmidt HG, Mamede S. The causes of errors in clinical reasoning: cognitive biases, knowledge deficits, and dual process thinking. *Acad Med* 2017 Jan;92(1):23-30. [doi: [10.1097/ACM.0000000000001421](https://doi.org/10.1097/ACM.0000000000001421)] [Medline: [27782919](https://pubmed.ncbi.nlm.nih.gov/27782919/)]
5. Seaman JB, Cohen TR, White DB. Reducing the stress on clinicians working in the ICU. *JAMA* 2018 Nov 20;320(19):1981-1982. [doi: [10.1001/jama.2018.14285](https://doi.org/10.1001/jama.2018.14285)] [Medline: [30357376](https://pubmed.ncbi.nlm.nih.gov/30357376/)]
6. Ahmadi O, Azizkhani R, Basravi M. Correlation between workplace and occupational burnout syndrome in nurses. *Adv Biomed Res* 2014;3(1):44 [FREE Full text] [doi: [10.4103/2277-9175.125751](https://doi.org/10.4103/2277-9175.125751)] [Medline: [24627852](https://pubmed.ncbi.nlm.nih.gov/24627852/)]
7. Pronovost P, Berenholtz S, Dorman T, Lipsett PA, Simmonds T, Haraden C. Improving communication in the ICU using daily goals. *J Crit Care* 2003 Jun;18(2):71-75. [doi: [10.1053/jcrc.2003.50008](https://doi.org/10.1053/jcrc.2003.50008)] [Medline: [12800116](https://pubmed.ncbi.nlm.nih.gov/12800116/)]
8. Williams M, Hevelone N, Alban RF, Hardy JP, Oxman DA, Garcia E, et al. Measuring communication in the surgical ICU: better communication equals better care. *J Am Coll Surg* 2010 Jan;210(1):17-22. [doi: [10.1016/j.jamcollsurg.2009.09.025](https://doi.org/10.1016/j.jamcollsurg.2009.09.025)] [Medline: [20123326](https://pubmed.ncbi.nlm.nih.gov/20123326/)]
9. Hsiang S, Allen D, Annan-Phan S, Bell K, Bolliger I, Chong T, et al. The effect of large-scale anti-contagion policies on the COVID-19 pandemic. *Nature* 2020 Aug 08;584(7820):262-267. [doi: [10.1038/s41586-020-2404-8](https://doi.org/10.1038/s41586-020-2404-8)] [Medline: [32512578](https://pubmed.ncbi.nlm.nih.gov/32512578/)]
10. Nicola M, Alsafi Z, Sohrabi C, Kerwan A, Al-Jabir A, Iosifidis C, et al. The socio-economic implications of the coronavirus pandemic (COVID-19): a review. *Int J Surg* 2020 Jun;78:185-193 [FREE Full text] [doi: [10.1016/j.ijsu.2020.04.018](https://doi.org/10.1016/j.ijsu.2020.04.018)] [Medline: [32305533](https://pubmed.ncbi.nlm.nih.gov/32305533/)]
11. Kamerlin S, Kasson P. Managing coronavirus disease 2019 spread with voluntary public health measures: Sweden as a case study for pandemic control. *Clin Infect Dis* 2020 Dec 15;71(12):3174-3181 [FREE Full text] [doi: [10.1093/cid/ciaa864](https://doi.org/10.1093/cid/ciaa864)] [Medline: [32609825](https://pubmed.ncbi.nlm.nih.gov/32609825/)]
12. Nandy K, Salunke A, Pathak SK, Pandey A, Doctor C, Puj K, et al. Coronavirus disease (COVID-19): a systematic review and meta-analysis to evaluate the impact of various comorbidities on serious events. *Diabetes Metab Syndr* 2020 Sep;14(5):1017-1025 [FREE Full text] [doi: [10.1016/j.dsx.2020.06.064](https://doi.org/10.1016/j.dsx.2020.06.064)] [Medline: [32634716](https://pubmed.ncbi.nlm.nih.gov/32634716/)]
13. Bushell V, Thomas L, Combes J. Inside the O2: the NHS Nightingale Hospital London education center. *J Interprof Care* 2020 Sep 29;34(5):698-701. [doi: [10.1080/13561820.2020.1823949](https://doi.org/10.1080/13561820.2020.1823949)] [Medline: [32990127](https://pubmed.ncbi.nlm.nih.gov/32990127/)]
14. Maassen O, Fritsch S, Gantner J, Deffge S, Kunze J, Marx G, et al. Future mobile device usage, requirements, and expectations of physicians in German university hospitals: web-based survey. *J Med Internet Res* 2020 Dec 21;22(12):e23955 [FREE Full text] [doi: [10.2196/23955](https://doi.org/10.2196/23955)] [Medline: [33346735](https://pubmed.ncbi.nlm.nih.gov/33346735/)]

15. Cheng CKY, Ip DKM, Cowling BJ, Ho LM, Leung GM, Lau EHY. Digital dashboard design using multiple data streams for disease surveillance with influenza surveillance as an example. *J Med Internet Res* 2011 Oct 14;13(4):e85 [FREE Full text] [doi: [10.2196/jmir.1658](https://doi.org/10.2196/jmir.1658)] [Medline: [22001082](https://pubmed.ncbi.nlm.nih.gov/22001082/)]
16. Thoma B, Bandi V, Carey R, Mondal D, Woods R, Martin L, et al. Developing a dashboard to meet Competence Committee needs: a design-based research project. *Can Med Educ J* 2020 Mar 07;11(1):e16-e34 [FREE Full text] [doi: [10.36834/cmej.68903](https://doi.org/10.36834/cmej.68903)] [Medline: [32215140](https://pubmed.ncbi.nlm.nih.gov/32215140/)]
17. Wright MC, Borbolla D, Waller RG, Del Fiol G, Reese T, Nesbitt P, et al. Critical care information display approaches and design frameworks: a systematic review and meta-analysis. *J Biomed Inform X* 2019 Sep;3:100041 [FREE Full text] [doi: [10.1016/j.yjbinx.2019.100041](https://doi.org/10.1016/j.yjbinx.2019.100041)] [Medline: [31423485](https://pubmed.ncbi.nlm.nih.gov/31423485/)]
18. Görges M, Staggers N. Evaluations of physiological monitoring displays: a systematic review. *J Clin Monit Comput* 2008 Feb 7;22(1):45-66. [doi: [10.1007/s10877-007-9106-8](https://doi.org/10.1007/s10877-007-9106-8)] [Medline: [18064532](https://pubmed.ncbi.nlm.nih.gov/18064532/)]
19. Olchanski N, Dziadzko MA, Tiong IC, Daniels CE, Peters SG, O'Horo JC, et al. Can a novel ICU data display positively affect patient outcomes and save lives? *J Med Syst* 2017 Sep 18;41(11):171. [doi: [10.1007/s10916-017-0810-8](https://doi.org/10.1007/s10916-017-0810-8)] [Medline: [28921446](https://pubmed.ncbi.nlm.nih.gov/28921446/)]
20. Ghazisaeidi M, Safdari R, Torabi M, Mirzaee M, Farzi J, Goodini A. Development of performance dashboards in healthcare sector: key practical issues. *Acta Inform Med* 2015 Oct;23(5):317-321 [FREE Full text] [doi: [10.5455/aim.2015.23.317-321](https://doi.org/10.5455/aim.2015.23.317-321)] [Medline: [26635442](https://pubmed.ncbi.nlm.nih.gov/26635442/)]
21. Ibrain J, Salluh F. A COVID-19 dashboard: data analytics to aid resource allocation in intensive care units. *ICU Manag Pract* 2020;20(1):37 [FREE Full text]
22. Johns G. The essential impact of context on organizational behavior. *Acad Manag Rev* 2006 Apr;31(2):386-408. [doi: [10.5465/amr.2006.20208687](https://doi.org/10.5465/amr.2006.20208687)]
23. Teixeira L, Ferreira C, Santos B. Using task analysis to improve the requirements elicitation in health information system. 2007 Presented at: 29th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC 2007); August 23-26, 2007; Lyon, France p. 3669. [doi: [10.1109/iembs.2007.4353127](https://doi.org/10.1109/iembs.2007.4353127)]
24. Martin JL, Clark DJ, Morgan SP, Crowe JA, Murphy E. A user-centred approach to requirements elicitation in medical device development: a case study from an industry perspective. *Appl Ergon* 2012 Jan;43(1):184-190 [FREE Full text] [doi: [10.1016/j.apergo.2011.05.002](https://doi.org/10.1016/j.apergo.2011.05.002)] [Medline: [21636072](https://pubmed.ncbi.nlm.nih.gov/21636072/)]
25. Preparing the healthcare workforce to deliver the digital future: An independent report on behalf of the Secretary of State for Health and Social Care. National Health Service. 2019. URL: <https://topol.hee.nhs.uk/wp-content/uploads/HEE-Topol-Review-2019.pdf> [accessed 2022-04-06]
26. Davey B, R. Parker K. Requirements elicitation problems: a literature analysis. *Issues Informing Sci Inf Technol* 2015;12:071-082. [doi: [10.28945/2211](https://doi.org/10.28945/2211)]
27. Liberati EG, Ruggiero F, Galuppo L, Gorli M, González-Lorenzo M, Maraldi M, et al. What hinders the uptake of computerized decision support systems in hospitals? A qualitative study and framework for implementation. *Implement Sci* 2017 Sep 15;12(1):113 [FREE Full text] [doi: [10.1186/s13012-017-0644-2](https://doi.org/10.1186/s13012-017-0644-2)] [Medline: [28915822](https://pubmed.ncbi.nlm.nih.gov/28915822/)]
28. Kahn JM. What we talk about when we talk about intensive care unit strain. *Annals ATS* 2014 Feb;11(2):219-220. [doi: [10.1513/annalsats.201311-406ed](https://doi.org/10.1513/annalsats.201311-406ed)]
29. Grubaugh ND, Hanage WP, Rasmussen AL. Making sense of mutation: what D614G means for the COVID-19 pandemic remains unclear. *Cell* 2020 Aug 20;182(4):794-795 [FREE Full text] [doi: [10.1016/j.cell.2020.06.040](https://doi.org/10.1016/j.cell.2020.06.040)] [Medline: [32697970](https://pubmed.ncbi.nlm.nih.gov/32697970/)]
30. Plante JA, Liu Y, Liu J, Xia H, Johnson BA, Lokugamage KG, et al. Spike mutation D614G alters SARS-CoV-2 fitness. *Nature* 2021 Apr 26;592(7852):116-121 [FREE Full text] [doi: [10.1038/s41586-020-2895-3](https://doi.org/10.1038/s41586-020-2895-3)] [Medline: [33106671](https://pubmed.ncbi.nlm.nih.gov/33106671/)]
31. Tiwari S, Rathore S. A methodology for the selection of requirement elicitation techniques. arXiv. 2017. URL: <http://arxiv.org/abs/1709.08481> [accessed 2022-04-06]
32. Kausar S, Tariq S, Riaz S, Khanum A. Guidelines for the selection of elicitation techniques. 2010 Presented at: 6th International Conference on Emerging Technologies (ICET); October 18-19, 2010; Islamabad, Pakistan p. 265-269. [doi: [10.1109/icet.2010.5638476](https://doi.org/10.1109/icet.2010.5638476)]
33. Maynard MT, Kennedy DM, Resick CJ. Teamwork in extreme environments: Lessons, challenges, and opportunities. *J Organ Behav* 2018 Jul 02;39(6):695-700. [doi: [10.1002/job.2302](https://doi.org/10.1002/job.2302)]
34. Qualitative Data Analysis Software. NVIVO. 2021. URL: <https://www.qsrinternational.com/nvivo-qualitative-data-analysis-software/home> [accessed 2022-04-06]
35. Braun V, Clarke V. Using thematic analysis in psychology. *Qual Res Psychol* 2006 Jan;3(2):77-101. [doi: [10.1191/1478088706qp063oa](https://doi.org/10.1191/1478088706qp063oa)]
36. Roberts K, Dowell A, Nie J. Attempting rigour and replicability in thematic analysis of qualitative research data; a case study of codebook development. *BMC Med Res Methodol* 2019 Mar 28;19(1):66 [FREE Full text] [doi: [10.1186/s12874-019-0707-y](https://doi.org/10.1186/s12874-019-0707-y)] [Medline: [30922220](https://pubmed.ncbi.nlm.nih.gov/30922220/)]
37. Fereday J, Muir-Cochrane E. Demonstrating rigor using thematic analysis: a hybrid approach of inductive and deductive coding and theme development. *Int J Qual Methods* 2016 Nov 29;5(1):80-92. [doi: [10.1177/160940690600500107](https://doi.org/10.1177/160940690600500107)]
38. Lingard L, Schryer CF, Spafford MM, Campbell SL. Negotiating the politics of identity in an interdisciplinary research team. *Qual Res* 2016 Aug 15;7(4):501-519. [doi: [10.1177/1468794107082305](https://doi.org/10.1177/1468794107082305)]

39. Greckhamer T, Koro-Ljungberg M, Cilesiz S, Hayes S. Demystifying interdisciplinary qualitative research. *Qual Inquiry* 2008 Mar 01;14(2):307-331. [doi: [10.1177/1077800407312049](https://doi.org/10.1177/1077800407312049)]
40. Frith H, Gleeson K. Clothing and embodiment: men managing body image and appearance. *Psychol Men Masculinity* 2004 Jan;5(1):40-48. [doi: [10.1037/1524-9220.5.1.40](https://doi.org/10.1037/1524-9220.5.1.40)]
41. Fuller TE, Pong DD, Piniella N, Pardo M, Bessa N, Yoon C, et al. Interactive digital health tools to engage patients and caregivers in discharge preparation: implementation study. *J Med Internet Res* 2020 Apr 28;22(4):e15573 [FREE Full text] [doi: [10.2196/15573](https://doi.org/10.2196/15573)] [Medline: [32343248](https://pubmed.ncbi.nlm.nih.gov/32343248/)]
42. Cosgriff CV, Ebner DK, Celi LA. Data sharing in the era of COVID-19. *Lancet Digit Health* 2020 May;2(5):e224 [FREE Full text] [doi: [10.1016/S2589-7500\(20\)30082-0](https://doi.org/10.1016/S2589-7500(20)30082-0)] [Medline: [32373785](https://pubmed.ncbi.nlm.nih.gov/32373785/)]
43. Bardram JE. The trouble with login: on usability and computer security in ubiquitous computing. *Pers Ubiquit Comput* 2005 Jul 23;9(6):357-367. [doi: [10.1007/s00779-005-0347-6](https://doi.org/10.1007/s00779-005-0347-6)]
44. Bourdeaux CP, Thomas MJ, Gould TH, Malhotra G, Jarvstad A, Jones T, et al. Increasing compliance with low tidal volume ventilation in the ICU with two nudge-based interventions: evaluation through intervention time-series analyses. *BMJ Open* 2016 May 26;6(5):e010129 [FREE Full text] [doi: [10.1136/bmjopen-2015-010129](https://doi.org/10.1136/bmjopen-2015-010129)] [Medline: [27230998](https://pubmed.ncbi.nlm.nih.gov/27230998/)]
45. Mitka M. Joint commission warns of alarm fatigue: multitude of alarms from monitoring devices problematic. *JAMA* 2013 Jun 12;309(22):2315-2316. [doi: [10.1001/jama.2013.6032](https://doi.org/10.1001/jama.2013.6032)] [Medline: [23757063](https://pubmed.ncbi.nlm.nih.gov/23757063/)]
46. Lewandowska K, Weisbrot M, Cieloszyk A, Mędrzycka-Dąbrowska W, Krupa S, Ozga D. Impact of alarm fatigue on the work of nurses in an intensive care environment-a systematic review. *Int J Environ Res Public Health* 2020 Nov 13;17(22):8409 [FREE Full text] [doi: [10.3390/ijerph17228409](https://doi.org/10.3390/ijerph17228409)] [Medline: [33202907](https://pubmed.ncbi.nlm.nih.gov/33202907/)]
47. Storm J, Chen HC. The relationships among alarm fatigue, compassion fatigue, burnout and compassion satisfaction in critical care and step-down nurses. *J Clin Nurs* 2021 Feb;30(3-4):443-453. [doi: [10.1111/jocn.15555](https://doi.org/10.1111/jocn.15555)] [Medline: [33174282](https://pubmed.ncbi.nlm.nih.gov/33174282/)]
48. Cobus V, Heuten W. To beep or not to beep? Evaluating modalities for multimodal ICU alarms. *Multimodal Technol Interact* 2019 Mar 09;3(1):15. [doi: [10.3390/mti3010015](https://doi.org/10.3390/mti3010015)]
49. Jenkin A, Abelson-Mitchell N, Cooper S. Patient handover: time for a change? *Accid Emerg Nurs* 2007 Jul;15(3):141-147. [doi: [10.1016/j.aeen.2007.04.004](https://doi.org/10.1016/j.aeen.2007.04.004)] [Medline: [17618118](https://pubmed.ncbi.nlm.nih.gov/17618118/)]
50. Zhang S, Wu Q, van Velthoven MH, Chen L, Car J, Rudan I, et al. Smartphone versus pen-and-paper data collection of infant feeding practices in rural China. *J Med Internet Res* 2012 Sep 18;14(5):e119 [FREE Full text] [doi: [10.2196/jmir.2183](https://doi.org/10.2196/jmir.2183)] [Medline: [22989894](https://pubmed.ncbi.nlm.nih.gov/22989894/)]
51. Ruth CJ, Huey SL, Krisher JT, Fothergill A, Gannon BM, Jones CE, et al. An electronic data capture framework (ConnEDCt) for global and public health research: design and implementation. *J Med Internet Res* 2020 Aug 13;22(8):e18580 [FREE Full text] [doi: [10.2196/18580](https://doi.org/10.2196/18580)] [Medline: [32788154](https://pubmed.ncbi.nlm.nih.gov/32788154/)]
52. Sbaffi L, Walton J, Blenkinsopp J, Walton G. Information overload in emergency medicine physicians: a multisite case study exploring the causes, impact, and solutions in four North England National Health Service Trusts. *J Med Internet Res* 2020 Jul 27;22(7):e19126 [FREE Full text] [doi: [10.2196/19126](https://doi.org/10.2196/19126)] [Medline: [32716313](https://pubmed.ncbi.nlm.nih.gov/32716313/)]
53. Vincent J, Marshall JC, Namendys-Silva SA, François B, Martin-Loeches I, Lipman J, ICON investigators. Assessment of the worldwide burden of critical illness: the intensive care over nations (ICON) audit. *Lancet Respir Med* 2014 May;2(5):380-386. [doi: [10.1016/S2213-2600\(14\)70061-X](https://doi.org/10.1016/S2213-2600(14)70061-X)] [Medline: [24740011](https://pubmed.ncbi.nlm.nih.gov/24740011/)]

Abbreviations

EHR: electronic health record
HLH: hemophagocytic lymphohistiocytosis
ICU: intensive care unit
MIMIC: Medical Information Mart for Intensive Care
NHS: National Health Service
PF: arterial oxygen to inspired oxygen ratio.
PPE: personal protective equipment
SOFA: Sequential Organ Failure Assessment

Edited by A Kushniruk; submitted 19.05.21; peer-reviewed by A Fruhling, T Aslanidis, M Wright; comments to author 08.07.21; revised version received 27.09.21; accepted 03.01.22; published 13.04.22.

Please cite as:

Davidson B, Ferrer Portillo KM, Wac M, McWilliams C, Bourdeaux C, Craddock I

Requirements for a Bespoke Intensive Care Unit Dashboard in Response to the COVID-19 Pandemic: Semistructured Interview Study
JMIR Hum Factors 2022;9(2):e30523

URL: <https://humanfactors.jmir.org/2022/2/e30523>

doi: [10.2196/30523](https://doi.org/10.2196/30523)

PMID: [35038301](https://pubmed.ncbi.nlm.nih.gov/35038301/)

©Brittany Davidson, Katuska Mara Ferrer Portillo, Marcell Wac, Chris McWilliams, Christopher Bourdeaux, Ian Craddock. Originally published in JMIR Human Factors (<https://humanfactors.jmir.org>), 13.04.2022. This is an open-access article distributed under the terms of the Creative Commons Attribution License (<https://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Human Factors, is properly cited. The complete bibliographic information, a link to the original publication on <https://humanfactors.jmir.org>, as well as this copyright and license information must be included.

Original Paper

Enhancing Clinical Information Display to Improve Patient Encounters: Human-Centered Design and Evaluation of the Parkinson Disease-BRIDGE Platform

Ethan G Brown¹, MD; Erica Schleimer¹, BSc; Ian O Bledsoe¹, MD, MS; William Rowles¹, BA; Nicolette A Miller¹, BSc; Stephan J Sanders², BMBS, PhD; Katherine P Rankin¹, PhD; Jill L Ostrem¹, MD; Caroline M Tanner^{1,3}, MD, PhD; Riley Bove¹, MD, MMSc

¹University of California San Francisco Weill Institute for Neurosciences, University of California San Francisco, San Francisco, CA, United States

²Department of Psychiatry, University of California San Francisco, San Francisco, CA, United States

³Parkinson Disease Research, Education, and Clinical Center, San Francisco Veterans Affairs Medical Center, San Francisco, CA, United States

Corresponding Author:

Riley Bove, MD, MMSc

University of California San Francisco Weill Institute for Neurosciences

University of California San Francisco

San Francisco, CA

United States

Phone: 1 415 353 2311

Email: riley.bove@ucsf.edu

Abstract

Background: People with Parkinson disease (PD) have a variety of complex medical problems that require detailed review at each clinical encounter for appropriate management. Care of other complex conditions has benefited from digital health solutions that efficiently integrate disparate clinical information. Although various digital approaches have been developed for research and care in PD, no digital solution to personalize and improve communication in a clinical encounter is readily available.

Objective: We intend to improve the efficacy and efficiency of clinical encounters with people with PD through the development of a platform (PD-BRIDGE) with personalized clinical information from the electronic health record (EHR) and patient-reported outcome (PRO) data.

Methods: Using human-centered design (HCD) processes, we engaged clinician and patient stakeholders in developing PD-BRIDGE through three phases: an inspiration phase involving focus groups and discussions with people having PD, an ideation phase generating preliminary mock-ups for feedback, and an implementation phase testing the platform. To qualitatively evaluate the platform, movement disorders neurologists and people with PD were sent questionnaires asking about the technical validity, usability, and clinical relevance of PD-BRIDGE after their encounter.

Results: The HCD process led to a platform with 4 modules. Among these, 3 modules that pulled data from the EHR include a longitudinal module showing motor ratings over time, a display module showing the most recently collected clinical rating scales, and another display module showing relevant laboratory values and diagnoses; the fourth module displays motor symptom fluctuation based on an at-home diary. In the implementation phase, PD-BRIDGE was used in 17 clinical encounters for patients cared for by 1 of 11 movement disorders neurologists. Most patients felt that PD-BRIDGE facilitated communication with their clinician (n=14, 83%) and helped them understand their disease trajectory (n=11, 65%) and their clinician's recommendations (n=11, 65%). Neurologists felt that PD-BRIDGE improved their ability to understand the patients' disease course (n=13, 75% of encounters), supported clinical care recommendations (n=15, 87%), and helped them communicate with their patients (n=14, 81%). In terms of improvements, neurologists noted that data in PD-BRIDGE were not exhaustive in 62% (n=11) of the encounters.

Conclusions: Integrating clinically relevant information from EHR and PRO data into a visually efficient platform (PD-BRIDGE) can facilitate clinical encounters with people with PD. Developing new modules with more disparate information could improve these complex encounters even further.

(*JMIR Hum Factors* 2022;9(2):e33967) doi:[10.2196/33967](https://doi.org/10.2196/33967)

KEYWORDS

human-centered design; personal health record; visualization in eHealth; Parkinson disease; digital health

Introduction

Although numerous digital health instruments have been developed, technology has thus far underdelivered when it comes to synthesizing clinical information in a coherent and an efficient way so that a clinician can use the same at the point of care (POC). As people with Parkinson disease (PD) experience problems in many different clinical domains—including motor, autonomic, cognitive, and sleep difficulties [1]—clinicians are faced with the challenge of soliciting and managing a broad range of symptoms. Medication management in people with PD is often highly personalized and based on symptoms that can change frequently, requiring detailed discussions in the clinic. Methods to quantify and track symptoms, whether through in-home mobile apps [2] or more granular quantitative metrics [3], have gained traction in PD and will present even more data streams requiring integration at the POC. Successful incorporation of this information into the clinical workflow, integrated with the already overwhelming amount of information available in the electronic health record (EHR) [4,5], is a critical challenge to overcome if clinicians are to deliver personalized and efficient care.

Researchers have addressed similar issues for other clinical conditions through platforms and dashboards that synthesize information from the EHR. Visualization dashboards have been used to manage the multiple information streams in intensive care units, where they reduce cognitive load [6] and improve quality metric adherence [7], which is variable in neurology [8]. The time required for an inpatient neurology consultation shortened after implementation of a clinical data review platform that integrated clinical information with vital signs, imaging results, and lab findings [9]. Recently, a clinician- and patient-facing platform was designed for multiple sclerosis to serve as a personalized visual aid for understanding the disease course [10]. A framework has also been proposed for integrating questionnaires administered outside usual clinical workflow to supplement the information in the EHR [11]. However, no standard tool with these capabilities is available that is tailored to the complex issues that arise in PD.

To meet this need, we developed a dashboard (PD-BRIDGE) that could be launched directly from the EHR to facilitate clinical interactions with people with PD. Our goal is to improve the efficiency and efficacy of clinical encounters with patients with PD. These encounters are highly complex because the wide-ranging symptoms people with PD experience demand multiple types of information to be considered for management, and this substantial time is needed to solicit and document these types of information. Any tool with this focus must be designed considering the intended users (ie, clinicians or patients). Such an instrument would be most effectively developed using human-centered design (HCD). HCD, when applied to digital health technology, describes a process that starts with identifying the needs of all stakeholders involved in the system that the technology hopes to change, continues with iterative feedback, and finally accounts for how the outcomes of the digital

intervention compare with the intended goals [12]. HCD seeks to reduce the reluctance and delay with which many health technologies are adopted into daily practice [13].

Therefore, we used HCD to design and develop a platform that integrates various data types from the EHR, as well as patient-reported outcomes (PROs), into 1 coherent dashboard that could be easily reviewed by physicians and patients as part of clinical workflows and clinically actionable. Here, we describe the HCD process resulting in the development of PD-BRIDGE according to proposed reporting guidelines for health research involving design [14]. We also report initial user experience.

Methods

Clinical and Research Setting

All study activities took place at the University of California, San Francisco (UCSF) Movement Disorders and Neuromodulation Center, a tertiary academic referral center with a busy clinical practice involving people with PD and other movement disorders. Our research team included clinicians with specialization in movement disorders (EB, IB, JO, and CT), clinicians with experience in digital health applications (RB, KR, and SS), a software engineer (ES), and a participant coordinator and data analyst (WR). These roles were chosen to guarantee familiarity with the challenges of clinical encounters with people with PD and the scope of how digital health solutions could address these challenges. PROs were completed remotely in the patients' homes prior to each visit, and some clinical consultations involved telemedicine.

Design Process

Phase I: Inspiration Phase

The first phase of HCD is focused on understanding the problem and empathizing with the users. PD-BRIDGE was adapted from BRIDGE, an established platform launched directly from the EHR to pull clinical data in real time to be actionable during the clinical encounter. For adaptability, BRIDGE was designed as a Substitutable Medical Applications and Reusable Technology on Fast Health Interoperability Resources (SMART on FHIR) [15,16] interface, which provides the technical capability for implementation in different EHR vendors and usage of different applications. BRIDGE converts multiple information streams (clinic-specific flow sheets, imaging data, patient questionnaires, and other EHR data elements) into visualization modules, or “widgets,” that can be customized accordingly to fit the needs of users with various clinical conditions, including multiple sclerosis [10] and other neurologic and neuropsychiatric diagnoses. Therefore, our first step was to identify the most useful modules to develop for PD-BRIDGE.

We began our inspiration phase with 2 focus groups, 1 with 14 movement disorders experts, as well as 5 individual discussions with people with PD. In this phase, we identified challenges

associated with systematically collecting and visualizing data, understanding patient histories, and gathering and understanding patients' daily patterns for medication adherence and side effects. The inspiration phase occurred over the course of 3 months.

Phase II: Ideation Phase

The ideation phase of HCD is focused on creating solutions for the problems defined in the inspiration phase. Through the focus groups, we generated mock-ups for preferred data visualization modules that could support the clinical information most often discussed and required for making decisions in clinical encounters with PD patients. Preliminary mock-up designs of these ideas were constructed and then presented to 5 physician/patient stakeholders for more feedback before implementation. Feedback was centered around summarizing complex histories, visualizing elements that could support medication management, and better visualization of longitudinal progression. The modules were then programmed and implemented on a live platform and made available to a select group of testers. The ideation phase occurred over the course of 6 months.

Phase III: Implementation Phase

The implementation phase is focused on building, testing, and iterating the solution. Once the PD-BRIDGE dashboard was built, user testing involved 2 stages. In the first stage, 2 movement disorders physicians—selected for their clinical volume and enthusiasm for digital technologies—agreed to use the platform for their upcoming regularly scheduled clinical encounters and to display it to their patients as appropriate. The clinicians then provided feedback regarding any potential software bugs, accuracy related to the medical chart, and requirement of further coding (eg, whether to code and display different formulations of a medication as 1 or 2 separate medications). Patients in these encounters were invited to fill out PROs before their scheduled visit and provide qualitative feedback regarding the usability and relevance of the tool after their clinical encounter. Further programmatic development then occurred to address qualitative feedback arising during this stage. This first stage took place over 3 months.

In the second stage, all 14 movement disorders physicians in our center were invited to use the platform in their upcoming regularly scheduled clinical encounters, with a goal of 30 total encounters. Each participating clinician signed an informed consent. Clinicians would identify patients having complex symptoms amenable to PD-BRIDGE and notify the research coordinator so they could be contacted in advance of the visit over telephone. After obtaining informed consent, the research

coordinator would obtain demographic information, basic information about the disease state, and then send the PROs to the participants.

During the clinical encounter, clinicians launched the PD-BRIDGE platform; both sets of users were then invited to complete user experience surveys (see [Multimedia Appendices 1-2](#)). Survey questions were developed based on the Health Information Technology Usability Evaluation Model [17]. This model integrates multiple usability theories including the Technology Acceptance Model (TAM) [18] and evaluates subjective (eg, perceived usefulness, perceived ease of use) and objective (eg, efficiency, effectiveness) outcomes. Specifically, patient surveys focused on the usefulness of PD-BRIDGE, the quality of communication throughout the encounter, comfort with the implementation and perceived security of the platform, and overall satisfaction of the visit. Clinician surveys focused on completeness of the data available, usability of PD-BRIDGE, the ability of the platform to facilitate the encounter, and overall satisfaction with the visit. Both surveys asked for specific feedback about visualizations and data in PD-BRIDGE. All surveys were in English and were administered using the REDCap platform; REDCap was also used to store survey responses. This second stage of implementation occurred over the course of 6 months, leading to a total design process timeline of 18 months.

Data Analysis

We constructed descriptive tables to summarize the demographic and clinical information about the participant cohort. We then summarized survey response data from clinicians and participant respondents. All descriptive analyses were performed using R (version 4.0.2; R Foundation for Statistical Computing). Tables were generated using the `gtsummary` package [19] and graphs were constructed using the `ggplot2` package in R [20].

Approvals and Consent

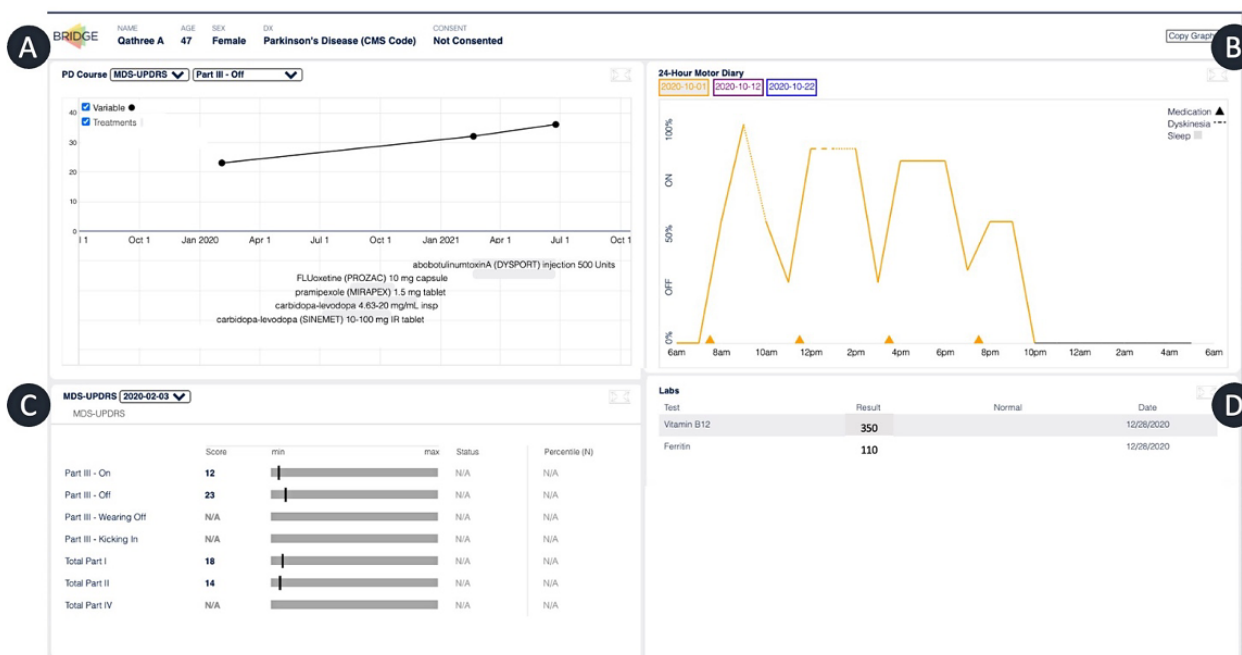
Evaluation of the platform and responses to the questionnaires were approved by the UCSF Institutional Review Board (IRB # 18-26148). Electronically signed informed consent was obtained from clinicians and patients prior to their participation.

Results

PD-Specific BRIDGE Modules Designed and Developed Through the HCD Process

Through stakeholder discussions, several modules were designed that would provide meaningful information during a clinical encounter with a person with PD ([Figure 1](#)).

Figure 1. Mock-up of the Parkinson disease-specific BRIDGE platform with (A) longitudinal measurement of motor scales, (B) visualization of a patient-entered motor diary, (C) cross sectional clinical scores, and (D) prior diagnoses and laboratory studies relevant to Parkinson disease. HCC: hierarchical condition category.



Adaptation of Existing Modules

To provide historical context to patients and clinicians, information about PD-related medication use and graphical display of the Movement Disorders Society Unified Parkinson’s Disease Rating Scale scores were determined to be highly relevant. These visualizations could be adapted from existing BRIDGE modules, with PD-specific data pulled from the EHR and displayed longitudinally (Figure 1A) as well as cross-sectionally in more detail (Figure 1B). Additionally, laboratory studies and comorbid conditions that are relevant to symptoms and medication management in PD were pulled from the EHR (Figure 1D).

New Module

Stakeholders agreed that a module displaying changes in motor symptoms over the course of the day based on a prospectively collected diary would facilitate clinical decision-making around medication changes and represent an advancement over patient/caregiver recall alone. Having reviewed the existing literature and iOS/Android app stores for such a feature and finding none, we designed a specific module to visualize data from a motor diary prospectively collected as an electronic PRO over the course of 24 hours prior to a patient’s visit. Our survey was initially based on the original paper diary for PD designed by Hauser et al [21], asking participants to define whether they were in the medication “ON” or “OFF” state, or had dyskinesia involving abnormal involuntary movements that occur as a complication of medications for PD that were either troublesome or not troublesome. However, patients noted trouble with the binary choice between ON and OFF and requested more granularity for choosing to what extent their medications were working. Therefore, we adapted the diary to allow for a graded response (see Multimedia Appendix 3), adjusted the visualization to express medication effects over the course of

the day from a scale of 0% to 100%, and overlaid them with indicators of the presence of dyskinesia and medication timing (Figure 1B). Thus, the ON time for a given participant could be interpreted as the area under the curve and used to justify and discuss changes in medication timing or dosage.

Modules Prioritized for a Future Round of Development

Additional modules were deemed desirable by some participants but were postponed to a future round of development after the initial testing phase. These modules either had a lower overall priority according to stakeholders or required a greater technical “lift.” The themes of these proposed modules include understanding longitudinal changes in nonmotor symptoms of PD (including mood and cognition), visualizing scores of neuropsychiatric testing, integrating neuroimaging and clinical videos, and visualizing changes in deep brain stimulation settings. PD-BRIDGE can also display quantitative motor and nonmotor ratings in relation to averages across the clinic or compared to age-adjusted normative values. Clinicians felt that these features would not be appropriate because advanced patients could be discouraged by seeing their information in this context, scores could be subjective, and normative values were not well established for PD. Hence, we did not include these features in this iteration but asked participants for their opinions in this area.

Implementation Phase: Patient Experience

In total, 34 patients consented to test the PD-BRIDGE platform in their clinical encounter and 32 filled out the motor diary. Of these, 21 completed the user experience survey after their visit (Table 1). Demographic and clinical characteristics of those who completed the survey were not significantly different from those who did not (data not shown). Overall, PD-BRIDGE was used in at least 17 clinical encounters; in 1 additional encounter,

a clinician reported using PD-BRIDGE whereas a patient did not.

Of the 17 patients reporting that their physician used PD-BRIDGE during their clinical encounter, 14 (83%) felt that PD-BRIDGE facilitated communication with their clinician, 11 (65%) thought it helped them understand their disease trajectory, and 11 (65%) felt that it helped them understand their clinician's

recommendations (Figure 2). Participants felt comfortable visualizing their own data points, and though not currently a feature of the PD-specific BRIDGE, 9 (50%) stated that they felt comfortable comparing their data with others' deidentified data, and 10 (56%) felt comfortable having their deidentified data used to inform decision-making for others. Importantly, no participants expressed concerns about the privacy of their data on the PD-BRIDGE platform (Figure 3).

Table 1. Characteristics of patient users of PD-BRIDGE (N=17).

Characteristic	Value
Age at visit in years, mean (SD)	66 (11)
Males, n (%)	11 (65)
Not Hispanic or Latino, n (%)	17 (100)
Race, n (%)	
Other	1 (5.9)
Unknown	1 (5.9)
White	15 (88)
Disease duration in years, mean (SD)	6.8 (4.0)

Figure 2. Responses from patients regarding the helpfulness of using PD-BRIDGE.

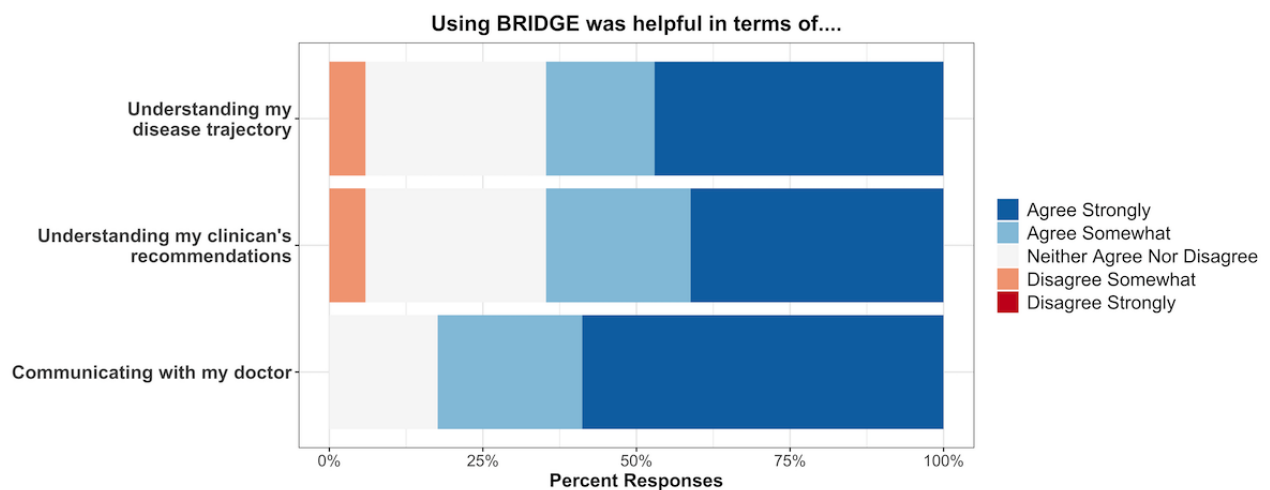
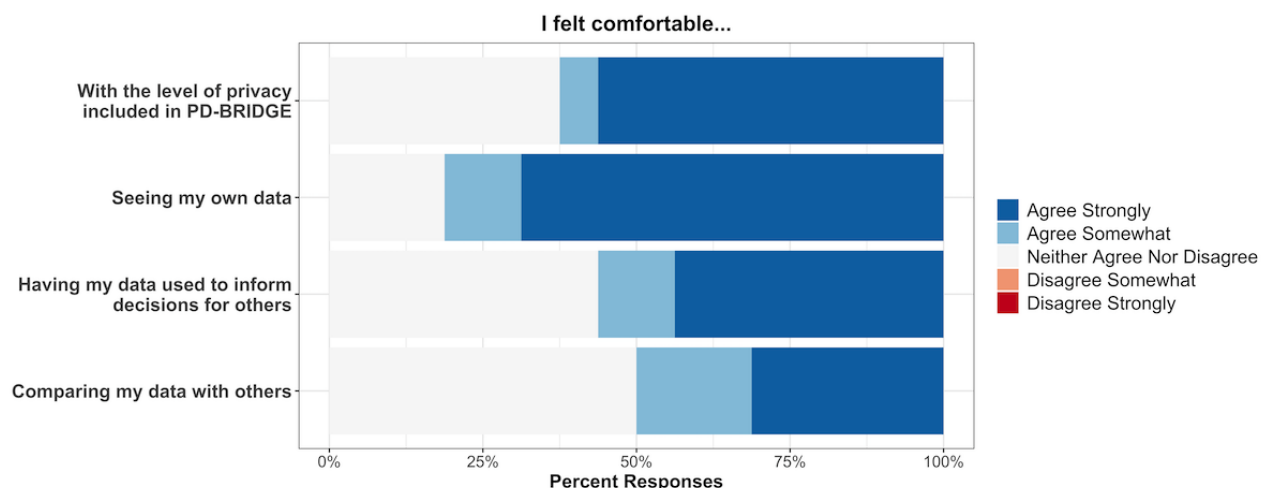


Figure 3. Responses from patients pertaining to their comfort with different aspects of PD-BRIDGE.



Specific comments were solicited from patient participants about what did and did not work well about the PD-BRIDGE platform. With respect to ease of use, 3 participants remarked that filling in data was fast and easy, whereas 1 found filling out the motor diary challenging. With respect to the likeability and usefulness of the visualizations, 5 appreciated the visualizations. Qualitative comments included feelings that PD-BRIDGE provided “a clearer picture of how (medications were) working at different times throughout the day/night,” that the “graph was helpful in explaining...symptoms during the day,” and that it showed “visual progression.” One participant specifically noted that discussing the visualizations facilitated “adjustment of...medication as it relates to wearing-off.” Another patient remarked that the platform did not yet include enough data to be helpful.

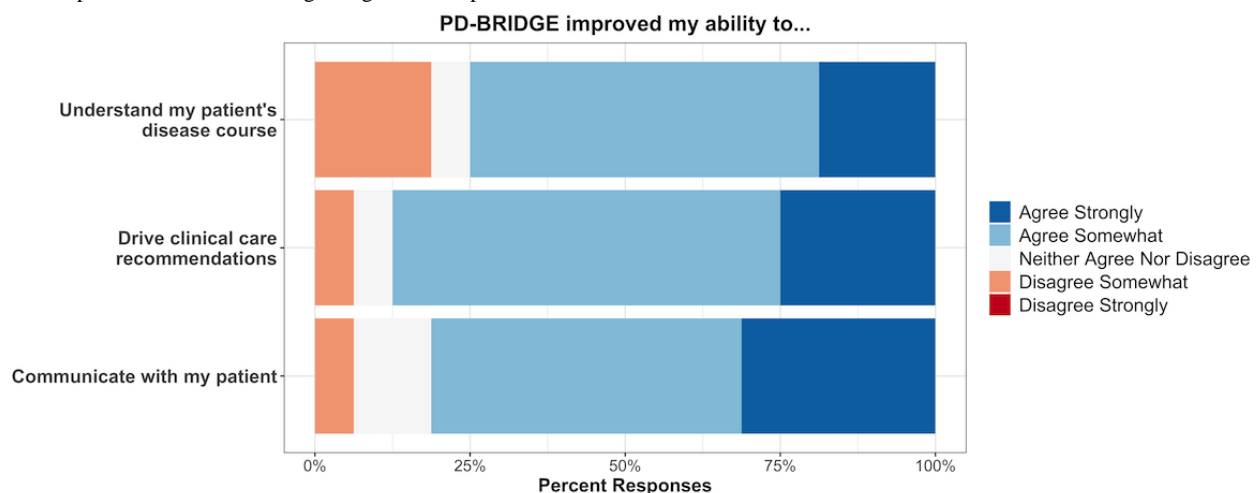
Implementation Phase: Neurologist Experience

In total, 11 movement disorders neurologists filled out a survey for a patient encounter, for a total of 16 encounters where PD-BRIDGE was used. With respect to data and visualizations, neurologists generally felt that they had the correct visual aids to explain their patients' disease to them (agreeing for $n=11$, 69% of encounters) and disagreed that they had difficulty communicating with their patient about their disease course ($n=14$, 88%) or their recommended care ($n=15$, 93%). Neurologists felt that PD-BRIDGE helped them “get on the

same page” as their patient 94% ($n=15$) of the time. Further, although neurologists felt that the data were up to date in 94% ($n=15$) of the encounters, they felt the data were not exhaustive in 62% ($n=10$) and that the EHR was more complete in 50% ($n=8$) of encounters. Still, during 44% ($n=7$) of the encounters, neurologists felt that the data in PD-BRIDGE were more complete than what could be retrieved through the EHR. In terms of improvements, 4 neurologists suggested that more data be imported into PD-BRIDGE to make it applicable to a wider range of patients, whereas others requested more features such as visualization of nonmotor symptom progression (1 neurologist) or uploading of clinical videos (1 neurologist).

With respect to perceived usefulness, neurologists felt that PD-BRIDGE helped with many clinical aspects of the office visit in the majority of the 16 encounters, including understanding their patients' disease course ($n=12$, 75%), driving clinical care recommendations ($n=14$, 87%), and communicating with their patient ($n=13$, 81%), as shown in Figure 4. When asked about specific features that worked, they mentioned the benefit of “visualizing (that) fluctuations were greatly reduced,” “demonstrat(ing) the need for increasing medications,” and “hav(ing) a visual aid for patients and clinicians to reference and to guide discussion.” Neurologists appreciated seeing “the (clinical) trajectory longitudinally” and the “ability to incorporate the MDS-UPDRS.”

Figure 4. Responses from clinicians regarding clinical aspects of PD-BRIDGE.



Discussion

Principal Results

Using iterative design techniques and engaging feedback from intended patient and clinician users, we developed a disease-specific platform designed to facilitate clinical encounters for the care of people with PD.

Although information overload occurs throughout different areas of medicine, PD is a condition that would particularly benefit from integrated and efficient delivery of clinical information. Symptoms in PD change frequently over the course of a typical day and vary from patient to patient, making modules such as the motor diary helpful to not only capture hour-by-hour fluctuations in patient experience but also visualize

these reports efficiently. These symptoms will likely be captured by mobile or wearable devices in the near future, creating another information stream that needs to be effectively displayed in the EHR. Furthermore, PD management involves multiple types of medications and procedures (eg, deep brain stimulation and botulinum toxin), which can evolve and become increasingly complex over time, resulting in a large amount of information to review and consider at each clinic visit. Designing ways to rapidly ingest and use this data to inform clinical decisions and assist in counseling patients is imperative.

We used an HCD process to develop a SMART on FHIR application to approach this problem. We decided that the HCD process would most effectively solicit the varied perspectives involved in clinical care of people with PD with an instrument that could then be adopted quickly into clinical workflows. We

also felt that the SMART on FHIR platform and modular approach of BRIDGE allowed for the greatest adaptability to various needs and would guarantee eventual transferability to other contexts. The HCD process was indeed effective in engaging clinicians and patients in all stages of design and allowed us to incorporate clinical workflows into platform development. However, a major challenge with this process was how to prioritize various requests and perspectives; even within the same subspecialty, clinicians have various workflows and find different types of information useful, such that selecting the most important and most widely applicable visualizations was challenging. Future iterations of PD-BRIDGE will include a wider array of modules to meet these requests, and the modular approach will allow for further personalization. We also noted that stakeholder enthusiasm for aspects of PD-BRIDGE, expressed in focus groups, did not always translate to engagement, and incorporating pilot testing and implementation into the design phase was an essential aspect of HCD to identify the barriers to adoption of the platform.

PD-BRIDGE facilitates clinical encounters by integrating information from disparate sources (eg, home diaries, elements in the EHR) into easily visualized displays. The modular design of the platform allows for easy adaptation to various subspecialties. Our PD stakeholders were enthusiastic about this platform and readily identified several key features that would increase efficiency in clinical encounters. Some were prioritized for the current version, and others will be integrated into future versions. However, the perceived benefit of a digital health product is not sufficient for its adoption, and clinical workflows can be especially challenging processes to change. The likelihood that a digital health instrument such as PD-BRIDGE will be used can be evaluated using a proposed structured framework that considers technical validation, clinical validation, usability, and cost [22].

From a technical perspective, the information that PD-BRIDGE represented was accurate when compared to the current gold standard, namely the EHR. The main feedback from neurologists for the PD-BRIDGE platform was that data were not as complete as that in the EHR. Although the EHR has a wealth of clinically relevant information, much of it is buried in free text and is challenging to extract in an automated and a reliable fashion. To address these issues, we developed an EHR flowsheet in parallel where specific clinically relevant information is collected in a manner accessible to PD-BRIDGE for integration in future versions. Notably, our institution's EHR (EPIC) makes flow sheets available to other institutions' developers through EPIC App Orchard, meaning that they can be easily downloaded and used in other subspecialty clinics with the same EHR. Another important technical feature of the EHR is privacy, which BRIDGE maintains by launching directly from within the EHR firewall. Reassuringly, patients did not express any concerns with PD-BRIDGE threatening privacy of their data.

From a clinical perspective, the major goal of PD-BRIDGE was to facilitate discussions about symptom management, which can be complex in PD. In our discussions, movement disorders neurologists indicated that their patient counseling usually relies on purely verbal conversations with no visual aid, though reading material may be provided to a patient afterward. Some

neurologists did use illustrations to convey their messages, but these were usually not patient-specific. PD-BRIDGE transformed patient-specific data into visualizations that were rated as clinically useful by patients and providers; having a visual aid helped translate the patients' verbal description of their symptoms and improved their understanding of the purpose of medication changes. Some patients also noted that PD-BRIDGE facilitated longitudinal understanding of their condition, which is a challenging disease characteristic to grasp from the EHR. In initial focus groups, neurologists had expressed a concern about patients visualizing how their data compared to others or allowing their data to be even seen in the aggregate form. Although our version of PD-BRIDGE did not include these features, it was reassuring that the patients we surveyed indicated no specific concern around these issues, and such features may be worthwhile to include in the future.

PD-BRIDGE also demonstrated sufficient usability, though we identified areas for improvement. The majority of PD-BRIDGE data are pulled automatically from the medical record, and this therefore places minimal burden on the clinician users, who can access all these data with a click at the POC. However, data must be entered discretely to be used in visualizations, and as PD-BRIDGE begins to incorporate more data streams, clinicians may need to change how they enter clinical data. PD-BRIDGE also relies on patient-entered PROs, including the 24-hour motor diary visualizations, which did require time from patients. Some patients appreciated the opportunity to list their symptoms, and some found it burdensome and error-prone. Therefore, these aspects of PD-BRIDGE may be more amenable to some clinicians and patients and not to others, and understanding these opinions can inform future implementation efforts. Future advances, such as a voice input option in lieu of keyboarding, may improve accessibility for motor-impaired patients.

Limitations

Despite this encouraging preliminary feedback, the current study's limitations require that further testing be conducted. We were only able to survey a small number of patients and neurologists and may not have captured the full range of feedback on the instrument. The results of the survey may also have suffered from selection bias; although we included the majority of movement disorders neurologists in our division, the groups of patients willing to test PD-BRIDGE were possibly already enthusiastic about this type of technology, even if we did not observe significant differences between participants and nonparticipants. Future studies can evaluate how PD-BRIDGE improves clinician efficiency, such as by reducing clicks in the EHR, or improves health outcomes, such as by emphasizing important topics from the patient perspective or reminding physicians of quality metrics.

Conclusions

Overall, this study shows the usefulness of adapting a platform that exists within the EHR to subspecialty-specific use. Future versions of PD-BRIDGE will integrate more information streams, such as images, clinical videos, PROs that capture more aspects of PD, and patient devices. As the complexity and breadth of clinical care in PD increases, such a platform will be

essential to translate the wealth of information into actionable clinical care.

Acknowledgments

Funding for design and testing of PD-BRIDGE was provided by National Institutes of Health National Center for Advancing Clinical Sciences (NIH-NCATS) pilot award through a UCSF Research Allocation Program Grant. BRIDGE was supported by a Weill Innovation Fund Award.

Authors' Contributions

EGB assisted in the design of the platform, analysis of the data, and drafting of the manuscript. ES was responsible for the design, development and iteration of the platform, and revision of the manuscript. WR was responsible for the development and implementation of patient-reported outcomes and surveys and revision of the manuscript. NAM was responsible for the development and iteration of the platform and revision of the manuscript. IOB was responsible for study design, feedback and revision of the platform, and revision of the manuscript. KPR was responsible for the design and development of the parent platform and revision of the manuscript. SJS was responsible for the design and development of the parent platform and revision of the manuscript. JLO was responsible for the study design, feedback and revision of the platform, and revision of the manuscript. CMT was responsible for the study design, feedback and revision of the platform, and revision of the manuscript. RB was responsible for the overall study design, design of the platform, and revision of the manuscript.

Conflicts of Interest

EGB receives research support from the Michael J Fox Foundation and the Gateway Institute for Brain Research Inc. He receives an honorarium from New England Journal of Medicine Knowledge+ in his role as Neurology Section Editor. IOB serves on the advisory board for Accorda and Ipsen and works as a consultant for Ideo Inc and Humancraft. He has received honoraria from the American Academy of Neurology and the California Hand Society. RB is the recipient of a National Multiple Sclerosis Harry Weaver Award. She has received research support from the National Multiple Sclerosis Society, the National Science Foundation, the Hilton Foundation, the California Initiative to Advance Precision Medicine, the Sherak Foundation, Biogen, and Roche Genentech. She has also received personal compensation for consulting from Alexion, Biogen, EMD Serono, Novartis, Roche Genentech, and Sanofi Genzyme. SJS has consulted for and received research funding from BioMarin. The other authors report no conflict of interest.

Multimedia Appendix 1

Patient postvisit questionnaire.

[PDF File (Adobe PDF File), 57 KB - [humanfactors_v9i2e33967_app1.pdf](#)]

Multimedia Appendix 2

Clinician postvisit questionnaire.

[PDF File (Adobe PDF File), 85 KB - [humanfactors_v9i2e33967_app2.pdf](#)]

Multimedia Appendix 3

Parkinson disease motor diary example.

[PDF File (Adobe PDF File), 93 KB - [humanfactors_v9i2e33967_app3.pdf](#)]

References

1. Poewe W, Seppi K, Tanner CM, Halliday GM, Brundin P, Volkman J, et al. Parkinson disease. Nat Rev Dis Primers 2017 Mar;3:17013. [doi: [10.1038/nrdp.2017.13](#)]
2. Bot BM, Suver C, Neto EC, Kellen M, Klein A, Bare C, et al. The mPower study, Parkinson disease mobile data collected using ResearchKit. Sci Data 2016 Mar;3:160011 [FREE Full text] [doi: [10.1038/sdata.2016.11](#)] [Medline: [26938265](#)]
3. Hill EJ, Mangleburg CG, Alfradique-Dunham I, Ripperger B, Stillwell A, Saade H, et al. Quantitative mobility measures complement the MDS-UPDRS for characterization of Parkinson's disease heterogeneity. Parkinsonism Relat Disord 2021 Mar;84:105-111 [FREE Full text] [doi: [10.1016/j.parkreldis.2021.02.006](#)] [Medline: [33607526](#)]
4. Schulte F, Fry E. Death by 1,000 clicks: where electronic health records went wrong. Fortune. 2019 Mar. URL: <https://khn.org/news/death-by-a-thousand-clicks/> [accessed 2022-03-24]
5. Arndt BG, Beasley JW, Watkinson MD, Temte JL, Tuan WJ, Sinsky CA, et al. Tethered to the EHR: primary care physician workload assessment using EHR event log data and time-motion observations. Ann Fam Med 2017 Sep;15(5):419-426 [FREE Full text] [doi: [10.1370/afm.2121](#)] [Medline: [28893811](#)]

6. Khairat SS, Dukkupati A, Lauria HA, Bice T, Travers D, Carson SS. The impact of visualization dashboards on quality of care and clinician satisfaction: integrative literature review. *JMIR Hum Factors* 2018 May;5(2):e22 [FREE Full text] [doi: [10.2196/humanfactors.9328](https://doi.org/10.2196/humanfactors.9328)] [Medline: [29853440](https://pubmed.ncbi.nlm.nih.gov/29853440/)]
7. Dye ME, Pugh C, Sala C, Scott TA, Wallace T, Grubb PH, et al. Developing a unit-based quality improvement program in a large neonatal ICU. *Jt Comm J Qual Patient Saf* 2021 Oct;47(10):654-662. [doi: [10.1016/j.jcjq.2021.06.003](https://doi.org/10.1016/j.jcjq.2021.06.003)] [Medline: [34284954](https://pubmed.ncbi.nlm.nih.gov/34284954/)]
8. Wilson AM, Benish SM, McCarthy L, Romano JG, Lundgren KB, Byrne M, et al. Quality of neurologic care in the United States. *Neurology* 2021 Aug;97(7):e651-e659. [doi: [10.1212/WNL.00000000000012378](https://doi.org/10.1212/WNL.00000000000012378)]
9. Kummer BR, Willey JZ, Zelenetz MJ, Hu Y, Sengupta S, Elkind MSV, et al. Neurological dashboards and consultation turnaround time at an academic medical center. *Appl Clin Inform* 2019 Oct;10(5):849-858 [FREE Full text] [doi: [10.1055/s-0039-1698465](https://doi.org/10.1055/s-0039-1698465)] [Medline: [31694054](https://pubmed.ncbi.nlm.nih.gov/31694054/)]
10. Bove R, Bruce CA, Lunders CK, Pearce JR, Liu J, Schleimer E, et al. Electronic health record technology designed for the clinical encounter. *Neurol Clin Pract* 2020 Oct;11(4):318-326. [doi: [10.1212/CPJ.0000000000000986](https://doi.org/10.1212/CPJ.0000000000000986)]
11. Philipson RG, Wu AD, Curtis WC, Jablonsky DJ, Hegde JV, McCloskey SA, et al. A practical guide for navigating the design, build, and clinical integration of electronic patient-reported outcomes in the radiation oncology department. *Pract Radiat Oncol* 2021 Jul;11(4):e376-e383. [doi: [10.1016/j.prro.2020.12.007](https://doi.org/10.1016/j.prro.2020.12.007)] [Medline: [33460827](https://pubmed.ncbi.nlm.nih.gov/33460827/)]
12. Melles M, Albayrak A, Goossens R. Innovating health care: key characteristics of human-centered design. *Int J Qual Health Care* 2021 Jan;33(Supplement_1):37-44 [FREE Full text] [doi: [10.1093/intqhc/mzaa127](https://doi.org/10.1093/intqhc/mzaa127)] [Medline: [33068104](https://pubmed.ncbi.nlm.nih.gov/33068104/)]
13. Erwin K, Krishnan JA. Redesigning healthcare to fit with people. *BMJ* 2016 Aug;354:i4536. [doi: [10.1136/bmj.i4536](https://doi.org/10.1136/bmj.i4536)] [Medline: [27553644](https://pubmed.ncbi.nlm.nih.gov/27553644/)]
14. Bazzano AN, Yan SD, Martin J, Mulhern E, Brown E, LaFond A, et al. Improving the reporting of health research involving design: a proposed guideline. *BMJ Glob Health* 2020 Feb;5(2):e002248. [doi: [10.1136/bmjgh-2019-002248](https://doi.org/10.1136/bmjgh-2019-002248)]
15. Mandl KD, Kohane IS. No small change for the health information economy. *New England Journal of Medicine* 2009 Mar;360(13):1278-1281. [doi: [10.1056/NEJMp0900411](https://doi.org/10.1056/NEJMp0900411)] [Medline: [19321867](https://pubmed.ncbi.nlm.nih.gov/19321867/)]
16. Mandel JC, Kreda DA, Mandl KD, Kohane IS, Ramoni RB. SMART on FHIR: a standards-based, interoperable apps platform for electronic health records. *J Am Med Inform Assoc* 2016 Feb;23(5):899-908. [doi: [10.1093/jamia/ocv189](https://doi.org/10.1093/jamia/ocv189)] [Medline: [26911829](https://pubmed.ncbi.nlm.nih.gov/26911829/)]
17. Brown W, Yen PY, Rojas M, Schnall R. Assessment of the Health IT Usability Evaluation Model (Health-ITUEM) for evaluating mobile health (mHealth) technology. *J Biomed Inform* 2013 Dec;46(6):1080-1087 [FREE Full text] [doi: [10.1016/j.jbi.2013.08.001](https://doi.org/10.1016/j.jbi.2013.08.001)] [Medline: [23973872](https://pubmed.ncbi.nlm.nih.gov/23973872/)]
18. Nguyen M, Fujioka J, Wentlandt K, Onabajo N, Wong I, Bhatia RS, et al. Using the technology acceptance model to explore health provider and administrator perceptions of the usefulness and ease of using technology in palliative care. *BMC Palliat Care* 2020 Sep;19:138. [doi: [10.1186/s12904-020-00644-8](https://doi.org/10.1186/s12904-020-00644-8)]
19. Sjoberg D, Whiting K, Curry M, Lavery J, Larmarange J. Reproducible summary tables with the gtsummary package. *R Journal* 2021 Jun;13(1):570-594 [FREE Full text] [doi: [10.32614/rj-2021-053](https://doi.org/10.32614/rj-2021-053)]
20. Wickham H. *ggplot2: Elegant Graphics for Data Analysis*. New York: Springer-Verlag; 2016.
21. Hauser RA, Friedlander J, Zesiewicz TA, Adler CH, Seeberger LC, O'Brien CF, et al. A home diary to assess functional status in patients with Parkinson's disease with motor fluctuations and dyskinesia. *Clin Neuropharmacol* 2000 Mar;23(2):75-81. [doi: [10.1097/00002826-200003000-00003](https://doi.org/10.1097/00002826-200003000-00003)] [Medline: [10803796](https://pubmed.ncbi.nlm.nih.gov/10803796/)]
22. Mathews SC, McShea MJ, Hanley CL, Ravitz A, Labrique AB, Cohen AB. Digital health: a path to validation. *NPJ Digit Med* 2019 May;2:38. [doi: [10.1038/s41746-019-0111-3](https://doi.org/10.1038/s41746-019-0111-3)]

Abbreviations

EHR: electronic health record

HCD: human-centered design

PD: Parkinson disease

POC: point of care

PRO: patient-reported outcome

SMART on FHIR: Substitutable Medical Applications and Reusable Technology on Fast Health Interoperability Resources

Edited by A Kushniruk; submitted 06.10.21; peer-reviewed by I Gottgens, P Macharia; comments to author 25.11.21; revised version received 11.01.22; accepted 07.02.22; published 06.05.22.

Please cite as:

*Brown EG, Schleimer E, Bledsoe IO, Rowles W, Miller NA, Sanders SJ, Rankin KP, Ostrem JL, Tanner CM, Bove R
Enhancing Clinical Information Display to Improve Patient Encounters: Human-Centered Design and Evaluation of the Parkinson
Disease-BRIDGE Platform*

JMIR Hum Factors 2022;9(2):e33967

URL: <https://humanfactors.jmir.org/2022/2/e33967>

doi: [10.2196/33967](https://doi.org/10.2196/33967)

PMID: [35522472](https://pubmed.ncbi.nlm.nih.gov/35522472/)

©Ethan G Brown, Erica Schleimer, Ian O Bledsoe, William Rowles, Nicolette A Miller, Stephan J Sanders, Katherine P Rankin, Jill L Ostrem, Caroline M Tanner, Riley Bove. Originally published in JMIR Human Factors (<https://humanfactors.jmir.org>), 06.05.2022. This is an open-access article distributed under the terms of the Creative Commons Attribution License (<https://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Human Factors, is properly cited. The complete bibliographic information, a link to the original publication on <https://humanfactors.jmir.org>, as well as this copyright and license information must be included.

Original Paper

An Electronic Data Capture Tool for Data Collection During Public Health Emergencies: Development and Usability Study

Joan Brown¹, MBA, EdD[‡]; Manas Bhatnagar², MS; Hugh Gordon³, MD; Jared Goodner³, MS; J Perren Cobb², MD; Karen Lutrick⁴, PhD

¹Clinical Operations Business Intelligence, The Keck School of Medicine of the University of Southern California, Los Angeles, CA, United States

²Department of Surgery, The Keck School of Medicine of the University of Southern California, Los Angeles, CA, United States

³Akido Labs Inc, Los Angeles, CA, United States

⁴Department of Family and Community Medicine, University of Arizona, Tucson, AZ, United States

[‡]Discovery Program for Emergency Preparedness

Corresponding Author:

Joan Brown, MBA, EdD

Clinical Operations Business Intelligence

The Keck School of Medicine of the University of Southern California

1520 San Pablo St

Los Angeles, CA, 90033

United States

Phone: 1 310 245 8079

Email: joanbrown@gmail.com

Abstract

Background: The Discovery Critical Care Research Network Program for Resilience and Emergency Preparedness (Discovery PREP) partnered with a third-party technology vendor to design and implement an electronic data capture tool that addressed multisite data collection challenges during public health emergencies (PHE) in the United States. The basis of the work was to design an electronic data capture tool and to prospectively gather data on usability from bedside clinicians during national health system stress queries and influenza observational studies.

Objective: The aim of this paper is to describe the lessons learned in the design and implementation of a novel electronic data capture tool with the goal of significantly increasing the nation's capability to manage real-time data collection and analysis during PHE.

Methods: A multiyear and multiphase design approach was taken to create an electronic data capture tool, which was used to pilot rapid data capture during a simulated PHE. Following the pilot, the study team retrospectively assessed the feasibility of automating the data captured by the electronic data capture tool directly from the electronic health record. In addition to user feedback during semistructured interviews, the System Usability Scale (SUS) questionnaire was used as a basis to evaluate the usability and performance of the electronic data capture tool.

Results: Participants included Discovery PREP physicians, their local administrators, and data collectors from tertiary-level academic medical centers at 5 different institutions. User feedback indicated that the designed system had an intuitive user interface and could be used to automate study communication tasks making for more efficient management of multisite studies. SUS questionnaire results classified the system as highly usable (SUS score 82.5/100). Automation of 17 (61%) of the 28 variables in the influenza observational study was deemed feasible during the exploration of automated versus manual data abstraction. The creation and use of the Project Meridian electronic data capture tool identified 6 key design requirements for multisite data collection, including the need for the following: (1) scalability irrespective of the type of participant; (2) a common data set across sites; (3) automated back end administrative capability (eg, reminders and a self-service status board); (4) multimedia communication pathways (eg, email and SMS text messaging); (5) interoperability and integration with local site information technology infrastructure; and (6) natural language processing to extract nondiscrete data elements.

Conclusions: The use of the electronic data capture tool in multiple multisite Discovery PREP clinical studies proved the feasibility of using the novel, cloud-based platform in practice. The lessons learned from this effort can be used to inform the improvement of ongoing global multisite data collection efforts during the COVID-19 pandemic and transform current manual

data abstraction approaches into reliable, real time, and automated information exchange. Future research is needed to expand the ability to perform automated multisite data extraction during a PHE and beyond.

(*JMIR Hum Factors* 2022;9(2):e35032) doi:[10.2196/35032](https://doi.org/10.2196/35032)

KEYWORDS

clinical research design; disaster management; informatics; public health emergencies; electronic data capture; design tenet; public health emergency; electronic data; EDCT; real time data

Introduction

Knowledge sharing during public health emergencies (PHE) is critical to managing swift and appropriate responses by key decision makers. Moreover, clinical responsibilities are typically increased, and dedicated research personnel may be lacking during a PHE. Despite the call to action from the medical community placed on data sharing for effective response, there remains a lack of standard best practice on information exchange during PHE, with no widely available platform mechanism to facilitate data sharing [1-4]. The absence of standards and technology challenges the ability of clinicians to develop a unified treatment plan to confront patients exposed to the PHE at hand. This has been evident since 2001, when the US Public Health System was challenged with the threat of an Anthrax outbreak [5]. Disparate information sources and unclear jurisdiction across local, state, and federal agencies prevent accurate knowledge sharing and aligned recommendations from decision makers [5]. The lack of information during PHE is a global challenge, as demonstrated in the data collection efforts during the Zika virus epidemic, Ebola outbreak [6], and most recently the COVID-19 pandemic [7-9]. This has been exacerbated during the COVID-19 pandemic, where data are needed to guide treatment protocols, but data sharing across a global spectrum is nonexistent or delayed [9-12]. Global standards and a system that allows for real-time learning during public health crisis are critical to our health care community's ability to respond to PHE [7-9,13-18].

Optimal responses to PHE require data-driven approaches that allow for prospective and real-time clinical data collection and dissemination that overcome the various challenges in data quality [18]. The current systems suffer from inadequate infrastructure for multisite clinical data capture [8,16,19,20], delays in dissemination of data due to lack of technical capacity [21], a lack of tools to manage the quality of data [20], and the absence of simple and straightforward interfaces that do not add to clinical burden of data collection during PHE [18]. To mitigate the known barriers to data collection during PHE, the Discovery Critical Care Research Network Program for Resilience and Emergency Preparedness (Discovery PREP [17]) partnered with Akido Labs, a third-party technology vendor, to develop a platform known as Project Meridian, a tool designed

for data capture and dissemination during PHE. Discovery PREP's experience with current research data capture platforms during national health system stress tests, and other PHE tabletop exercises, indicated excessive person-hour effort required to coordinate data collection from multiple sites in a simulated PHE [22-25]. Thus, Discovery PREP began investigating novel methods toward multisite clinical data extraction with the goal of significantly increasing the nation's capability to manage real-time clinical data collection and analysis during PHE. Exploration proceeded with the design and development of a technology-agnostic electronic data capture tool that could facilitate multisite automated data extraction and storage. Following the development of the electronic data capture tool, the feasibility of advancing data capture using automated data extraction compared to manual data entry was assessed in 2 observational studies [26-29]. This paper describes the technical process and lessons learned from this effort, concluding with recommendations for improvement of data sharing platforms during PHE.

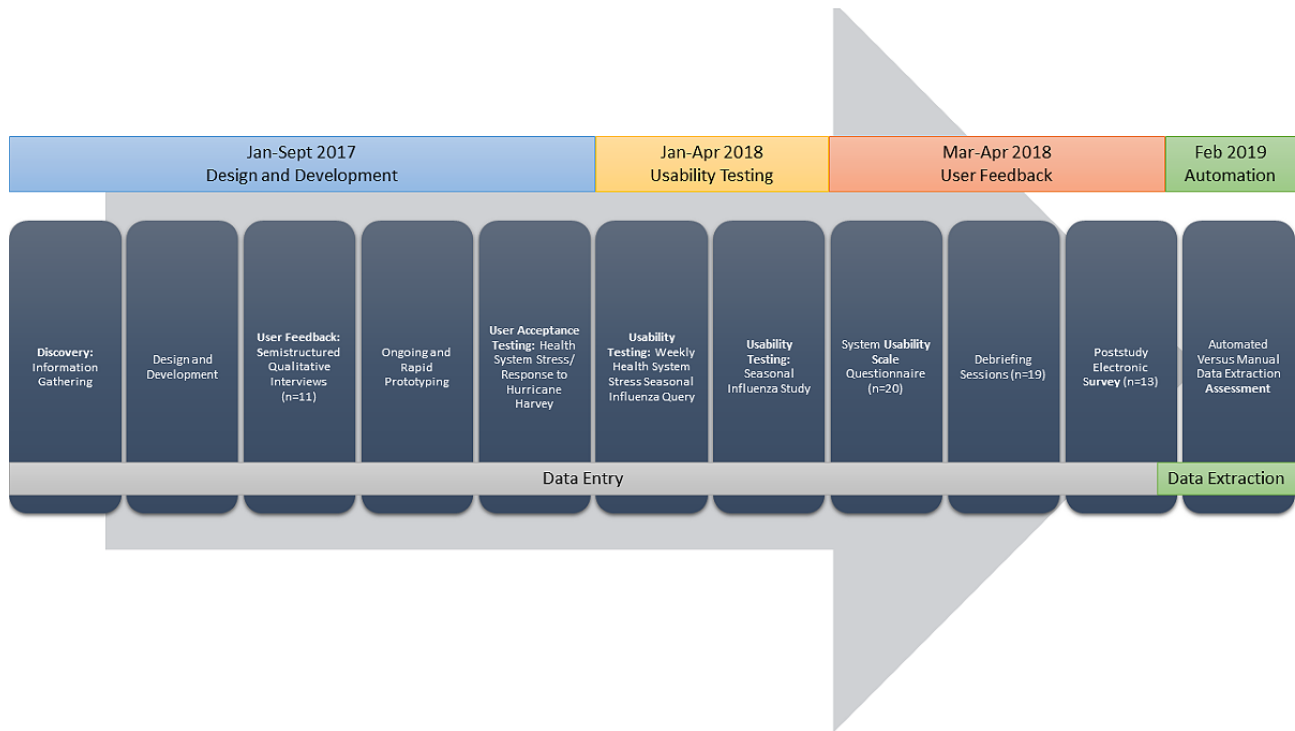
Methods

Overview

A multiyear and multiphase approach was taken to develop the electronic data capture tool as visualized in the design timeline (Figure 1). The tool was first developed and piloted for rapid data capture and then expanded to assess the feasibility of automated clinical data extraction. The design and evaluation of the electronic data capture tool spanned from January 2017 to April 2018.

Reducing the burden of data collection was a key design principle for the electronic data capture tool, as clinical responsibilities typically increase during PHE, and the availability of research personnel is insufficient to capture the volume of data need for robust clinical trials and their analysis, especially for the critically ill or injured [2,3,18]. The electronic data capture tool was designed with an intuitive data entry interface to reduce time and effort for data entry with the added capability to enter data on a smartphone. Ease of use was combined with considerations for scalability across multiple institutions to eliminate manual administration processes and bridge the gap created by disparate platforms.

Figure 1. Timeline.



Platform Design and Development

The participants included Discovery PREP physicians, their local administrators, and data collectors from tertiary-level academic medical centers at 5 geographically distributed institutions (University of Southern California, Washington University in St. Louis, Baylor University, Mayo Clinic, and Duke University). Design, development, rapid prototyping, and user feedback took place between January 2017 and July 2017. Information was gathered prior to the development of the

electronic data capture tool to identify the unmet needs and solidify design specifications. Individual semistructured interviews were conducted over a 45-minute duration. A total of 11 participants were interviewed, including 3 (27%) physician researchers, 3 (27%) data collectors, 4 (36%) administrators, and 1 (9%) biostatistician to understand workflows and data collection challenges during a PHE. The initial predevelopment interviews were qualitative in nature and elicited information on data collection processes and limitations of current electronic data capture tools (Table 1).

Table 1. Qualitative interview questions prior to electronic data capture tool development.

Question #	Question detail
1	When do you complete CRFs ^a with respect to enrollment time?
2	How are new study subjects identified?
3	How are new subjects communicated to data collector?
4	What is your process for collecting data for the CRF?
5	What are the pain points you experience with REDCap? ^b
6	Pain points with the last study you participated in?

^aCRF: case report form.

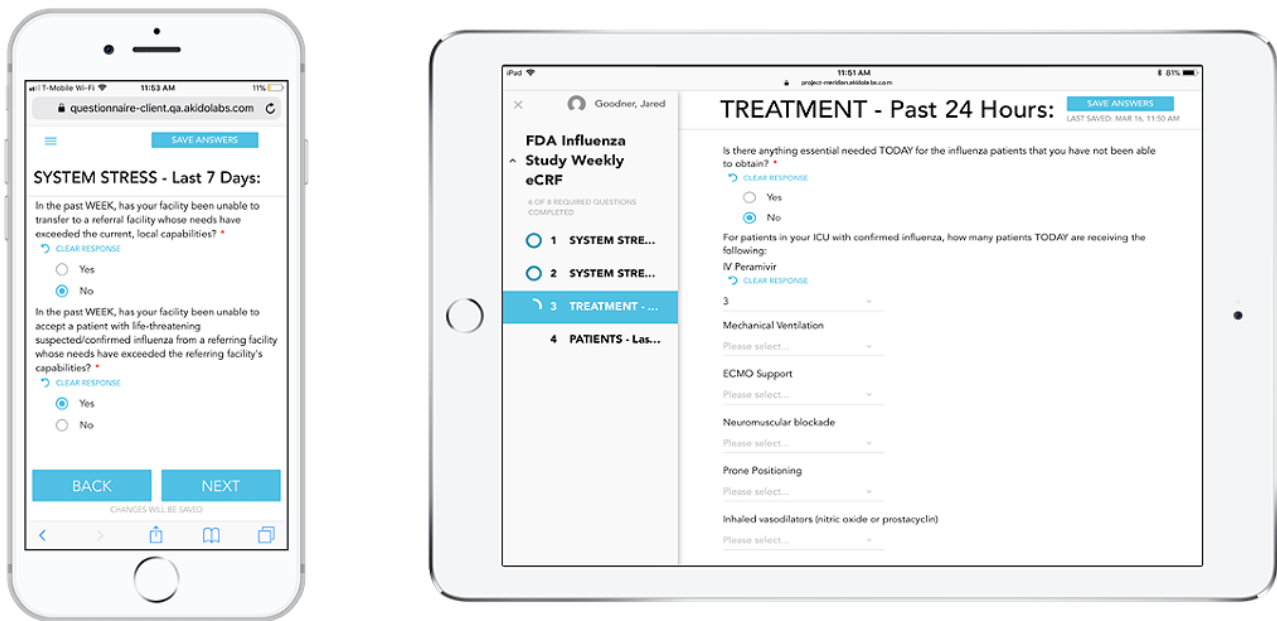
^bREDCap: Research Electronic Data Capture.

Project Meridian was designed to be powered by the Akido Labs Development Environment. The latter was designed to enable modern development in a health care environment by abstracting four core unique complexities specific to this industry, including security of patient health information, compliance, interoperability, and governance (Multimedia Appendix 1). User-centered design practices with an eye toward a simple user interface were the basis of the design of the user interface and prototypes (Figure 2; Multimedia Appendices 2 and 3). Postdevelopment, user prototyping interviews focused

on feedback regarding electronic data capture tool prototypes. An agile approach of rapid iteration following user feedback was taken to enhance the electronic data capture tool following each interview in preparation for the next, following a hypothesize-design-test learning loop. Interviews were performed both in person and remotely via screen sharing, as needed. The interviews were semistructured, and users conducted standardized tasks while observed by the investigator team including the following: Discovery PREP administrative team, Akido Labs engineers, and a notetaker to capture user

feedback on functionality, user experience, and messaging to guide usage. All clicks, mouse movements, and time required to accomplish specific tasks were recorded for analysis and used to refine the platform design.

Figure 2. Project Meridian mobile capability screenshots. ECMO: extracorporeal membrane oxygenation; eCRF: Electronic Case Report Form; FDA: Food and Drug Administration; ICU: intensive care unit; IV: intravenous.



Usability and Pilot Study Testing

Additional feedback was gathered during user acceptance testing (UAT). UAT was performed using the two following scenarios: (1) a Discovery PREP health system stress query over 401 participants in August 2017 (Figure 3), and (2) 34 Society of Critical Care Medicine participants affected by Hurricane Harvey in the state of Texas in September 2017 (Figure 4). The chart on the left for both Figures 3 and 4 show the breakdown of responder practice setting. The map in Figure 3 illustrates the map of responders superimposed on population density. The map in Figure 4 displays the number of responses to the Hurricane Harvey query. Additional feature enhancements were assessed based on the feedback gathered during UAT. Following

UAT, the electronic data capture tool was used to facilitate data collection for 2 clinical studies encompassing 403 users across the United States. Both clinical studies involved gathering information on the impact of seasonal influenza on health system stress. The studies were conducted with 12 sites for 17 weeks. The first study involved a predefined set of users with a large data collection form including 151 patient-level clinical data elements. The second encompassed a brief data collection form with 20 questions with census and health system stress level data. Data were collected weekly from health care systems using the Project Meridian platform. Following these studies, a subset of users (n=20) completed the System Usability Scale (SUS) questionnaire [30,31], 19 (95%) participated in debriefing sessions, and 13 (65%) completed a poststudy survey.

Figure 3. User acceptance testing (UAT) map of responders, National Health System Stress Query (n=401). ED: emergency department; ICU: intensive care unit.

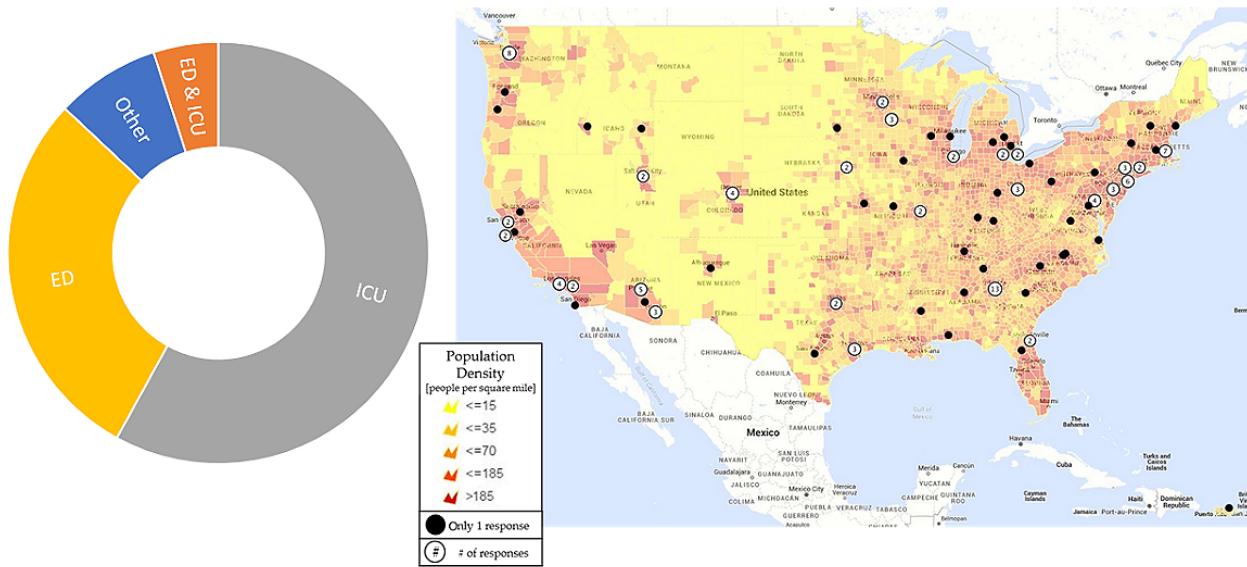
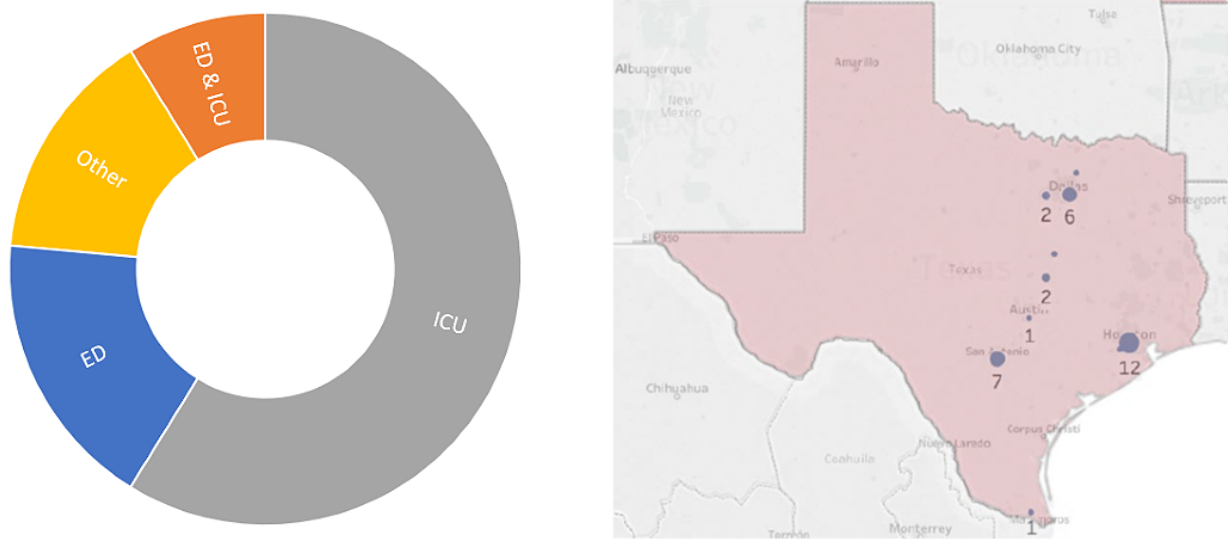


Figure 4. User acceptance testing (UAT) map of responders, Hurricane Harvey (n=20). ED: emergency department; ICU: intensive care unit.



Ethics Approval

Both studies were approved under the University of Southern California Investigational Review Board (HS-16-00948).

Automated Versus Manual Data Extraction

Seasonal influenza was used as a proxy for a PHE during the comparison of automated and manual data collection [29]. Patients in an intensive care unit (ICU) with laboratory-confirmed influenza were enrolled into an investigational review board–approved observational study (HS-17-00837). At a single institution, patient selection and data collection were completed using two methods in parallel: research personnel effort (manual) and querying of institutional clinical data warehouse (automated). Data were collected over a 2-week period using a consensus, previously reported tiered

case report form (CRF). Tier 1 of the CRF sought demographics, diagnoses, and lab results as well as supportive care details from the first 24 hours of ICU visit. Tier 2 sought more detailed clinical data from disease onset to patient discharge. Tier 1 was used for comparison in the feasibility test. The automated approach required the identification of relevant patients and gathering of key data elements by executing daily automated queries to an institutional clinical data warehouse. Data were stored and compared for accuracy following the 2-week period.

Results

Platform Design and Development

During the design phase, the results of the initial qualitative interviews highlighted the following themes: (1) the need to

automate data entry; (2) the need to automate frequent study communications and coordination tasks; (3) the importance of ease-of-access and usability; and (4) the need to enable real-time data reporting to stakeholders during a PHE. Identification of

these needs led to the inclusion of multiple feature enhancements within the Project Meridian platform prior to product launch (Textbox 1).

Textbox 1. Project Meridian feature enhancements.

Feature description

- Gamification—leaderboard for number of responses
- Text message-based survey initiation
- Refer a colleague (if primary responder not on clinical service, or new responder)
- Redesign of automated survey email (improving call to action)
- Improving visibility of case report form completion rate
- Common view for members of one study team (one institution). All case report forms visible to all study data collectors at a given institution
- Advanced query functions to prompt individuals or their sites

Usability and Pilot Study Testing

Design, development, and UAT of the platform occurred over a 9-month period. During usability testing, using the observational studies, data entry personnel reported increased awareness of data entry completeness with the use of site level summary dashboards. Additionally, Discovery PREP study administrators reported that the automation of scheduled personalized emails to the study participants reduced study administration time by an estimated 80% compared to previous studies. The results of the SUS questionnaire [30,31] classified the system among the 90th percentile of a broad class of systems evaluated [30] and was therefore highly usable (SUS score 82.5/100).

Automated Versus Manual Data Extraction

The automated and manual data extraction pilot for patient selection independently identified the correct patients (N=4) during the 2-week study period. Completion of Tier 1 of the CRF per patient was 100% (28/28) via manual approach and 61% (17/28) via automation. Compared with manually collected data, automated data were 50% (70/141) identical and 13% (18/141) different. Variables such as demographics, ventilator status, and availability of lab values were identical. The individual lab values pulled in the first 24 hours of ICU admission were not always identical as there were multiple values available for some patients within that first 24-hour period. Values for pregnancy status, preadmit events, coinfections, and means of identification were missing. Data obtained through automated means had an inherent delay of up to 24 hours due to the use of the data warehouse infrastructure. Manually collected data had an average delay of 2-days between fulfillment of inclusion criteria and enrollment into the study.

Discussion

The electronic data capture tool designed and tested proved highly usable and capable of collecting critical information during PHE test scenarios. One of the lessons learned globally during the recent COVID-19 pandemic is the importance of standardized real-time data collection, analysis, and reporting [7,32]. Prior to the pandemic, Discovery PREP investigators

and federal partners developed a novel data capture system to manage multisite data collection to address the all-hazards core data set used to characterize serious illness, injuries, and resource requirements during PHE [18]. The design and implementation of the Project Meridian electronic data capture tool was Discovery PREP's successful solution to enhance coordinated data collection capabilities during PHEs by addressing the pain points experienced by the clinical community during multisite data collection. Discovery PREP continued to leverage and report on the use of the Project Meridian platform in subsequent national studies [33,34].

Throughout this design and use process, Discovery PREP learned that specific design tenets need to be addressed to successfully gather essential information during a PHE. These tenets include the following:

1. Gathering data to assess a nationwide health system stress during influenza seasons involved collecting data from a bedside clinician (N=403) or an individual institution (N=12) [26-28]. Thus, the data collection system needs to be scalable and adaptable to the number and type of participants.
2. The data gathered during an event may include multiple types of case report forms with a combination of similar and differing variables that often require repeat measurements. For example, one of the observational studies was a weekly query to assess health system stress, while the other study was a single-report event with the same set of variables but with additional clinical content. Furthermore, a common data dictionary was created across Discovery PREP participating institutions to ensure the alignment of data collection across the sites. Thus, a data collection system must be able to accommodate a common consensus data set, with repeated measures across studies, and aggregate data for analysis and reporting to regional and federal government agencies.
3. Automating study administrative and communication tasks (eg, reminder emails) reduced the amount of manual administration for the study. Additionally, the status board (eg, leaderboard) served as a self-service visual to assess individual responses compared to others and to drive an

increase in participant response. Thus, a data collection system should automate communication tasks and incorporate a status board for self-service and to encourage participation, especially during PHEs such as the COVID-19 pandemic.

4. The participants noted that during a busy clinical shift, text messaging was a more effective way to obtain a rapid response. Thus, a data collection system needs to adapt to the preferred communication method of the participant, which may vary across time and institutions.
5. Automation of data and reduction of data acquisition time requires a highly interoperable system that integrates with the variety of platforms used at various institutions. Thus, a data collection system should provide the flexibility and functionality to integrate with local information technology infrastructure for automated and near-real time data capture.
6. In a single institution, the identification of eligible patients was reliably accomplished using automation. Additionally, 50% of the data collected manually for one of the observational studies was identically gathered through automation. However, when comparing the manual versus automated data extraction process, only discrete, categorical data fields were available. Text blocks within progress, operative, and discharge notes or the history and physical notes could not be automated for our purposes. Thus, an optimal data collection system should include natural language processing capability with access to these types of domains to fully automate local data extraction.

Extensive work is needed to meet the needs of rapid data collection during a PHE. This has been evident during the COVID-19 pandemic, where surveillance efforts have underlined the benefits of creating a Clinical Informatics Digital Hub for monitoring and for clinical trial data management [32]. To expand the findings in this report, more investigation is needed to assess the following: feasibility of real-time automation; the use of synchronization protocols as needed in areas challenged by unreliable or slow internet access [35]; the use of natural language processing to capture unstructured data [36]; and application of artificial intelligence to expand our ability to respond to a rapidly evolving disease [37]. With a lack of common regional or federal PHE reporting standards in the United States, third-party integration platforms such as Project Meridian can provide essential flexible infrastructure.

Rapid data collection is critical to an optimized national and international response [6,32,36]. Discovery PREP addressed this need by building and piloting an electronic data capture tool that was successful in collecting coordinated and real-time multisite data to assess health system stress and evaluated treatment protocols for seasonal influenza across the United States. The lessons learned from this report should be leveraged to improve data collection efforts and provide the foundation for further investigations focused on the evolution of manual data abstraction into reliable, real-time, and automated information exchange.

Acknowledgments

We would like to acknowledge the contribution of several Discovery Program for Emergency Preparedness partners including the participating academic sites (University of Southern California, Washington University in St. Louis, Baylor University, Mayo Clinic, and Duke University), and the technical team from Akido Labs Inc.

The work herein was funded in part by contracts from the Food and Drug Administration and the Biomedical Advanced Research and Development Authority (FDA BAA-12-00118) and the Assistant Secretary for Preparedness and Response (ASPR), ASPR-BARDA HHSO 100201300016A.

Disclaimer

The contents of this publication are the sole responsibility of the authors and do not necessarily reflect the views, opinions, or policies of any government entity. Mention of trade names, commercial products, or organizations does not imply endorsement by the US government.

Some of the information described herein was presented in abstract form at the 48th Critical Care Congress of the Society of Critical Care Medicine, February 17-20, 2019, San Diego, California, United States, and the 49th Annual Critical Care Congress, February 16-19, 2020, Orlando, Florida, United States.

Conflicts of Interest

None declared.

Multimedia Appendix 1

Akido Labs Development Environment overview.

[PNG File, 484 KB - [humanfactors_v9i2e35032_app1.png](#)]

Multimedia Appendix 2

Project Meridian query status and leaderboard screenshots.

[PNG File, 158 KB - [humanfactors_v9i2e35032_app2.png](#)]

Multimedia Appendix 3

Project Meridian SMS text messaging and email communication screenshots.

[\[PNG File , 298 KB - humanfactors_v9i2e35032_app3.png \]](#)**References**

1. Chretien J, Rivers CM, Johansson MA. Make Data Sharing Routine to Prepare for Public Health Emergencies. *PLoS Med* 2016 Aug 16;13(8):e1002109 [FREE Full text] [doi: [10.1371/journal.pmed.1002109](https://doi.org/10.1371/journal.pmed.1002109)] [Medline: [27529422](https://pubmed.ncbi.nlm.nih.gov/27529422/)]
2. Delaunay S, Kahn P, Tatay M, Liu J. Knowledge sharing during public health emergencies: from global call to effective implementation. *Bull. World Health Organ* 2016 Apr 01;94(4):236-236A. [doi: [10.2471/blt.16.172650](https://doi.org/10.2471/blt.16.172650)]
3. Lurie N, Manolio T, Patterson AP, Collins F, Frieden T. Research as a Part of Public Health Emergency Response. *N Engl J Med* 2013 Mar 28;368(13):1251-1255. [doi: [10.1056/nejmsb1209510](https://doi.org/10.1056/nejmsb1209510)]
4. Whitty CJM, Mundel T, Farrar J, Heymann DL, Davies SC, Walport MJ. Providing incentives to share data early in health emergencies: the role of journal editors. *The Lancet* 2015 Nov;386(10006):1797-1798. [doi: [10.1016/s0140-6736\(15\)00758-8](https://doi.org/10.1016/s0140-6736(15)00758-8)]
5. Gursky E, Inglesby T, O'Toole T. Anthrax 2001: observations on the medical and public health response. *Biosecur Bioterror* 2003;1(2):97-110. [doi: [10.1089/153871303766275763](https://doi.org/10.1089/153871303766275763)] [Medline: [15040187](https://pubmed.ncbi.nlm.nih.gov/15040187/)]
6. Kratochvil CJ, Evans L, Ribner BS, Lowe JJ, Harvey MC, Hunt RC, et al. The National Ebola Training and Education Center: Preparing the United States for Ebola and Other Special Pathogens. *Health Secur* 2017 Jun;15(3):253-260 [FREE Full text] [doi: [10.1089/hs.2017.0005](https://doi.org/10.1089/hs.2017.0005)] [Medline: [28636442](https://pubmed.ncbi.nlm.nih.gov/28636442/)]
7. Akhvediani T, Ali SM, Angus DC, Arabi YM, Ashraf S, Baillie JK, et al. Global outbreak research: harmony not hegemony. *The Lancet Infectious Diseases* 2020 Jul;20(7):770-772. [doi: [10.1016/S1473-3099\(20\)30440-0](https://doi.org/10.1016/S1473-3099(20)30440-0)]
8. Foraker RE, Lai AM, Kannampallil TG, Woeltje KF, Trolard AM, Payne PRO. Transmission dynamics: Data sharing in the COVID-19 era. *Learn Health Syst* 2020 Jun 28:e10235 [FREE Full text] [doi: [10.1002/lrh2.10235](https://doi.org/10.1002/lrh2.10235)] [Medline: [32838037](https://pubmed.ncbi.nlm.nih.gov/32838037/)]
9. Cosgriff CV, Ebner DK, Celi LA. Data sharing in the era of COVID-19. *The Lancet Digital Health* 2020 May;2(5):e224. [doi: [10.1016/s2589-7500\(20\)30082-0](https://doi.org/10.1016/s2589-7500(20)30082-0)]
10. Dron L, Dillman A, Zoratti MJ, Haggstrom J, Mills EJ, Park JJH. Clinical Trial Data Sharing for COVID-19-Related Research. *J Med Internet Res* 2021 Mar 12;23(3):e26718 [FREE Full text] [doi: [10.2196/26718](https://doi.org/10.2196/26718)] [Medline: [33684053](https://pubmed.ncbi.nlm.nih.gov/33684053/)]
11. Ossom-Williamson P, Williams IM, Kim K, Kindratt TB. Reporting and Availability of COVID-19 Demographic Data by US Health Departments (April to October 2020): Observational Study. *JMIR Public Health Surveill* 2021 Apr 06;7(4):e24288 [FREE Full text] [doi: [10.2196/24288](https://doi.org/10.2196/24288)] [Medline: [33821804](https://pubmed.ncbi.nlm.nih.gov/33821804/)]
12. Gates B. Responding to Covid-19 — A Once-in-a-Century Pandemic? *N Engl J Med* 2020 Apr 30;382(18):1677-1679. [doi: [10.1056/nejmp2003762](https://doi.org/10.1056/nejmp2003762)]
13. Littler K, Boon W, Carson G, Depoortere E, Mathewson S, Mietchen D, et al. Progress in promoting data sharing in public health emergencies. *Bull. World Health Organ* 2017 Apr 01;95(4):243-243. [doi: [10.2471/blt.17.192096](https://doi.org/10.2471/blt.17.192096)]
14. Moorthy V, Henao Restrepo AM, Preziosi M, Swaminathan S. Data sharing for novel coronavirus (COVID-19). *Bull. World Health Organ* 2020 Mar 01;98(3):150-150. [doi: [10.2471/blt.20.251561](https://doi.org/10.2471/blt.20.251561)]
15. Peiffer-Smadja N, Maatoug R, Lescure F, D'Ortenzio E, Pineau J, King J. Machine Learning for COVID-19 needs global collaboration and data-sharing. *Nat Mach Intell* 2020 May 22;2(6):293-294. [doi: [10.1038/s42256-020-0181-6](https://doi.org/10.1038/s42256-020-0181-6)]
16. Brown J, Bhatnagar M, Gordon H, Lutrick K, Goodner J, Blum J, et al. Clinical Data Extraction During Public Health Emergencies: A Blockchain Technology Assessment. *Biomed Instrum Technol* 2021 Jul 01;55(3):103-111. [doi: [10.2345/0890-8205-55.3.103](https://doi.org/10.2345/0890-8205-55.3.103)] [Medline: [34460906](https://pubmed.ncbi.nlm.nih.gov/34460906/)]
17. Cobb JP. Clinical Investigation During Public Health Emergencies: The Resilience Intelligence Network. *Am J Public Health* 2019 Sep;109(S4):S268-S270. [doi: [10.2105/ajph.2019.305215](https://doi.org/10.2105/ajph.2019.305215)]
18. Murphy DJ, Rubinson L, Blum J, Isakov A, Bhagwanjee S, Cairns CB, et al. Development of a Core Clinical Dataset to Characterize Serious Illness, Injuries, and Resource Requirements for Acute Medical Responses to Public Health Emergencies. *Critical Care Medicine* 2015;43(11):2403-2408. [doi: [10.1097/ccm.0000000000001274](https://doi.org/10.1097/ccm.0000000000001274)]
19. Dye C, Bartolomeos K, Moorthy V, Kieny MP. Data sharing in public health emergencies: a call to researchers. *Bull. World Health Organ* 2016 Mar 01;94(3):158. [doi: [10.2471/blt.16.170860](https://doi.org/10.2471/blt.16.170860)]
20. White E, Collins Rossetti S, Karipineni N, Maviglia S, Bartz R, Bhagwanjee S, et al. Optimization of Data Collection during Public Health Emergencies—Experience with APACHE II Score. *ACI Open* 2019 Apr 10;03(01):e18-e25. [doi: [10.1055/s-0039-1684001](https://doi.org/10.1055/s-0039-1684001)]
21. Modjarrad K, Moorthy VS, Millett P, Gsell P, Roth C, Kieny M. Developing Global Norms for Sharing Data and Results during Public Health Emergencies. *PLoS Med* 2016 Jan 5;13(1):e1001935 [FREE Full text] [doi: [10.1371/journal.pmed.1001935](https://doi.org/10.1371/journal.pmed.1001935)] [Medline: [26731342](https://pubmed.ncbi.nlm.nih.gov/26731342/)]
22. Society of Critical Care Medicine, U.S. Department of Health and Human Services. 106: Longitudinal assessment of U.S. health system stress during influenza seasons: pulse program. *Critical Care Medicine* 2018;46(1):36-36. [doi: [10.1097/01.ccm.0000528126.27672.86](https://doi.org/10.1097/01.ccm.0000528126.27672.86)]
23. Society of Critical Care Medicine. 113: Electronic data capture during public health emergencies. *Critical Care Medicine* 2018;46(1):39. [doi: [10.1097/01.ccm.0000528133.03909.a9](https://doi.org/10.1097/01.ccm.0000528133.03909.a9)]

24. Society of Critical Care Medicine. 655: Variability in treatment of influenza in the critically ill: a 2017 cohort study. *Critical Care Medicine* 2018;46(1):314. [doi: [10.1097/01.ccm.0000528670.07598.ee](https://doi.org/10.1097/01.ccm.0000528670.07598.ee)]
25. Read T, White E, Karipineni N, Ignatov M, Shanmugam M, Rocha B, et al. Utilization of Cost Effective Tools for Queries on Healthcare System Stress. *AMIA* 2016:1564 [[FREE Full text](#)]
26. University of Southern California. 237: Real-time clinical data collection during public health emergencies: 48-hour feasibility study. *Critical Care Medicine* 2019;47(1):100. [doi: [10.1097/01.ccm.0000550992.57579.a3](https://doi.org/10.1097/01.ccm.0000550992.57579.a3)]
27. University of Southern California. 236: ICU resource limitations during peak seasonal influenza: results of a national feasibility study. *Critical Care Medicine* 2019;47(1):100. [doi: [10.1097/01.ccm.0000550991.66931.5e](https://doi.org/10.1097/01.ccm.0000550991.66931.5e)]
28. Lane C, Bhatnagar M, Lutrick K, Maves RC, Weiner D, Rios Olvera D, et al. ICU Resource Limitations During Peak Seasonal Influenza: Results of a 2018 National Feasibility Study. *Crit Care Explor* 2022 Jan;4(1):e0606 [[FREE Full text](#)] [doi: [10.1097/CCE.0000000000000606](https://doi.org/10.1097/CCE.0000000000000606)] [Medline: [35018345](https://pubmed.ncbi.nlm.nih.gov/35018345/)]
29. University of Southern California. 298: A pilot study of automated data querying for public health emergencies. *Critical care medicine* 2020;48(1):131. [doi: [10.1097/01.ccm.0000619548.77699.33](https://doi.org/10.1097/01.ccm.0000619548.77699.33)]
30. System Usability Scale (SUS). usability.gov. URL: <https://www.usability.gov/how-to-and-tools/methods/system-usability-scale.html> [accessed 2018-03-01]
31. Bangor A, Kortum P, Miller J. Determining what individual SUS scores mean: Adding an adjective rating scale. *Journal of usability studies* 2009;4(3):114-123 [[FREE Full text](#)]
32. de Lusignan S, Jones N, Dorward J, Byford R, Liyanage H, Briggs J, et al. The Oxford Royal College of General Practitioners Clinical Informatics Digital Hub: Protocol to Develop Extended COVID-19 Surveillance and Trial Platforms. *JMIR Public Health Surveill* 2020 Jul 02;6(3):e19773 [[FREE Full text](#)] [doi: [10.2196/19773](https://doi.org/10.2196/19773)] [Medline: [32484782](https://pubmed.ncbi.nlm.nih.gov/32484782/)]
33. University of Southern California. 113: Design of an electronic clinical data capture system for use in public health emergencies. *Critical Care Medicine* 2019;47(1):39. [doi: [10.1097/01.ccm.0000550870.74251.08](https://doi.org/10.1097/01.ccm.0000550870.74251.08)]
34. Lane C, Maves R, Badke C, Weiner D, Uyeki T, Perren CJ. 303: National, real-time monitoring of hospital response during the 2018-2019 influenza season. *Critical Care Medicine* 2020;48(1):134. [doi: [10.1097/01.ccm.0000619568.22480.1c](https://doi.org/10.1097/01.ccm.0000619568.22480.1c)]
35. Ruth CJ, Huey SL, Krisher JT, Fothergill A, Gannon BM, Jones CE, et al. An Electronic Data Capture Framework (ConnEDCt) for Global and Public Health Research: Design and Implementation. *J Med Internet Res* 2020 Aug 13;22(8):e18580 [[FREE Full text](#)] [doi: [10.2196/18580](https://doi.org/10.2196/18580)] [Medline: [32788154](https://pubmed.ncbi.nlm.nih.gov/32788154/)]
36. Neuraz A, Lerner I, Digan W, Paris N, Tsopra R, Rogier A, AP-HP/Universities/INSERM COVID-19 Research Collaboration; AP-HP COVID CDR Initiative. Natural Language Processing for Rapid Response to Emergent Diseases: Case Study of Calcium Channel Blockers and Hypertension in the COVID-19 Pandemic. *J Med Internet Res* 2020 Aug 14;22(8):e20773 [[FREE Full text](#)] [doi: [10.2196/20773](https://doi.org/10.2196/20773)] [Medline: [32759101](https://pubmed.ncbi.nlm.nih.gov/32759101/)]
37. Izquierdo JL, Ancochea J, Savana COVID-19 Research Group, Soriano JB. Clinical Characteristics and Prognostic Factors for Intensive Care Unit Admission of Patients With COVID-19: Retrospective Study Using Machine Learning and Natural Language Processing. *J Med Internet Res* 2020 Oct 28;22(10):e21801 [[FREE Full text](#)] [doi: [10.2196/21801](https://doi.org/10.2196/21801)] [Medline: [33090964](https://pubmed.ncbi.nlm.nih.gov/33090964/)]

Abbreviations

CRF: case report form

Discovery PREP: Discovery Critical Care Research Network Program for Resilience and Emergency Preparedness

ICU: intensive care unit

PHE: public health emergencies

SUS: System Usability Scale

UAT: user acceptance testing

Edited by A Kushniruk; submitted 17.11.21; peer-reviewed by C Weerth, J Sharp; comments to author 06.03.22; revised version received 23.04.22; accepted 23.04.22; published 09.06.22.

Please cite as:

Brown J, Bhatnagar M, Gordon H, Goodner J, Cobb JP, Lutrick K

An Electronic Data Capture Tool for Data Collection During Public Health Emergencies: Development and Usability Study

JMIR Hum Factors 2022;9(2):e35032

URL: <https://humanfactors.jmir.org/2022/2/e35032>

doi: [10.2196/35032](https://doi.org/10.2196/35032)

PMID: [35679114](https://pubmed.ncbi.nlm.nih.gov/35679114/)

©Joan Brown, Manas Bhatnagar, Hugh Gordon, Jared Goodner, J Perren Cobb, Karen Lutrick. Originally published in JMIR Human Factors (<https://humanfactors.jmir.org>), 09.06.2022. This is an open-access article distributed under the terms of the Creative Commons Attribution License (<https://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Human Factors, is properly cited. The complete bibliographic information, a link to the original publication on <https://humanfactors.jmir.org>, as well as this copyright and license information must be included.

Original Paper

Usability of the Swedish Accessible Electronic Health Record: Qualitative Survey Study

Maria Hägglund¹, MSc, PhD; Isabella Scandurra², MSc, PhD

¹Healthcare Sciences and e-Health, Department of Women's and Children's Health, Uppsala University, Uppsala, Sweden

²Centre of Empirical Research on Information Systems, School of Business, Örebro University, Örebro, Sweden

Corresponding Author:

Maria Hägglund, MSc, PhD

Healthcare Sciences and e-Health

Department of Women's and Children's Health

Uppsala University

Akademiska Sjukhuset

Dag Hammarskjölds väg 14B, 1tr

Uppsala, 751 85

Sweden

Phone: 46 729999381

Email: maria.hagglund@kbh.uu.se

Abstract

Background: Patient portals are increasingly being implemented worldwide to ensure that patients have timely access to their health data, including patients' access to their electronic health records. In Sweden, the e-service Journalen is a national patient-accessible electronic health record (PAEHR), accessible on the web through the national patient portal. User characteristics and perceived benefits of using a PAEHR will influence behavioral intentions to use and adoption; however, poor usability, which increases effort expectancy, may have a negative impact. Therefore, it is of interest to further explore how users of the PAEHR Journalen perceive its usability and usefulness.

Objective: On the basis of the analysis of the survey respondents' experiences of the usability of the Swedish PAEHR, this study aimed to identify specific usability problems that may need to be addressed in the future.

Methods: A survey study was conducted to elicit opinions and experiences of patients using Journalen. Data were collected from June to October 2016. The questionnaire included a free-text question regarding the usability of the system, and the responses were analyzed using content analysis with a sociotechnical framework as guidance when grouping identified usability issues.

Results: During the survey period, 423,141 users logged into Journalen, of whom 2587 (0.61%) completed the survey (unique users who logged in; response rate 0.61%). Of the 2587 respondents, 186 (7.19%) provided free-text comments on the usability questions. The analysis resulted in 19 categories, which could be grouped under 7 of the 8 dimensions in the sociotechnical framework of Sittig and Singh. The most frequently mentioned problems were related to regional access limitations, structure and navigation of the patient portal, and language and understanding.

Conclusions: Although the survey respondents, who were also end users of the PAEHR Journalen, were overall satisfied with its usability, they also experienced important challenges when accessing their records. For all patients to be able to reap the benefits of record access, it is essential to understand both the usability challenges they encounter and, more broadly, how policies, regulations, and technical implementation decisions affect the usefulness of record access. The results presented here are specific to the Swedish PAEHR Journalen but also provide important insights into how design and implementation of record access can be improved in any context.

(*JMIR Hum Factors* 2022;9(2):e37192) doi:[10.2196/37192](https://doi.org/10.2196/37192)

KEYWORDS

usability; evaluation; patient-accessible electronic health records; open notes; patient portals; mobile phone

Introduction

Background

Patient portals are increasingly being implemented worldwide to ensure that patients have timely access to their health data and a means of communicating with health care professionals and managing their health care [1]. Although they were originally mostly for administration, patients are now gaining access to not only test results but also their full electronic health record (EHR), including notes written by health care professionals.

Patient-accessible EHRs (PAEHRs) have been or are being implemented in many countries, such as Finland, France, Norway, Australia, Denmark, Canada, the United Kingdom, and Sweden [2]. In some countries, these are local implementations at a specific hospital or region, whereas others have national solutions [3]. Differences in strategies and approaches have affected uptake and impact, and in several countries, the implementation progress has been slow because of legal constraints [4,5] and concerns about, for example, security and privacy among medical professionals [6-8].

In the United States, the OpenNotes initiative for providing patients access to their EHR began as a pilot and evaluation project that included 105 volunteer primary care physicians and their 19,000 patients [9,10]. The initiative started in 2010 and has since spread throughout the United States [11]. As of April 2021, a federal rule requires US health care providers to allow patients access to all health information in their EHR [12,13].

In Sweden, *Journalen* is a national PAEHR that is accessible via the web through the national patient portal 1177.se [14]. The PAEHR service accesses EHR information from different EHR systems used throughout Swedish health care organizations through a national health information exchange platform [15,16]. Therefore, patients have one access point for all their health record information [14]. Since the first Swedish region began providing its inhabitants web-based access to their health records in 2012, all other regions have connected to the national infrastructure and the PAEHR *Journalen*, with the last connection in March 2018. Different regions have also made different choices regarding how much of their information would be made available to patients [14,17].

A challenge that is frequently described internationally is the low adoption rate of patient portals and PAEHRs. This is often attributed to either perceived low usefulness or poor usability in combination with low eHealth literacy among users [18]. Therefore, it is of interest to explore how users of PAEHR *Journalen* perceive its usability and usefulness.

In this study, we analyzed data on usability issues from a national survey conducted among patients who used the PAEHR *Journalen*. A first analysis of the main results from the survey was published in 2018 [17] and contained an overview of the full survey. Usability was assessed in the survey using the System Usability Scale (SUS) [19,20], and the results of the SUS analysis indicated that the PAEHR was rated fairly highly by the respondents (81 on the SUS scale) [21]. However, as such, the SUS scale does not give any indication of what types

of usability problems end users experience or the severity. Therefore, in this study, we analyzed qualitative free-text comments related to the usability of the PAEHR from the same national survey.

Objectives

On the basis of the analysis of the survey respondents' experiences of the usability of the Swedish PAEHR, this study aimed to identify specific usability problems that may need to be addressed in the future. Regardless of whether the problems were frequent or disturbed *all* users, or whether the usability problems were severe or important to *some* users, the aim was to cover different aspects and indicate where future efforts need to be put.

Methods

Overview

A survey study was conducted to elicit the opinions and experiences of patients using *Journalen*. Participants were recruited through the national PAEHR *Journalen*. When patients logged into *Journalen*, they received a request for voluntary survey participation together with information about the study.

At the time of data collection (June to October 2016), not all regions were providing patients access to their records through *Journalen*, and among those who did, the level of transparency varied [17]. Of 22 health care providers (21 regions and 1 private health care provider), 18 (82%) gave patients web-based access to notes in the record, whereas only 8 (36%) gave access to laboratory results and 7 (32%) to immunizations [17]. Notes from psychiatric care were shared by only 2 (9%) health care providers, a number that has increased rapidly since [22].

Data Collection

A questionnaire was designed covering different topic areas with 24 questions, in Swedish (see the full questionnaire in the study by Moll et al [17]), including questions regarding the usability of the system using SUS [21], followed by a free-text comment where the respondents could add anything they wanted regarding the usability of the system. Thus, the resulting free-text comments were the material for this study.

The usability and technical functionality of the electronic questionnaire were not tested before fielding the questionnaire. However, participants received information about whom to contact in case of technical issues.

The collected data were managed by the national eHealth service provider Inera AB in accordance with the security requirements presented in the ethical application and approved by the Regional Ethical Review Board. The survey data were stored in the same database system as the PAEHR, indicating that the collected data, including the patient ID, had the same security protection as all patient information handled in the PAEHR. The patient ID was stored during the collection period to ensure that patients did not leave duplicate responses. When the collection period was completed, the patient ID was removed, and all stored information was anonymized. The anonymized data set was exported to the researchers for analysis.

Data Analysis

Overall, 2587 patients from 21 county councils completed the survey. The number of respondents for each county council or region varied. Only completed questionnaires were included for analysis, as the answers were stored in the database only when the respondent chose to submit them on the last page.

In this study, we focused on free-text answers related to the usability of Journalen. Free-text answers regarding usability were analyzed through inductive content analysis, as proposed by Graneheim and Lundman [23]. Questions regarding demography and the perceived usefulness of the overall survey are also presented.

In total, 186 respondents voluntarily provided free-text answers to the question about Journalen's usability. The answers were read independently by both authors (MH and IS), and *meaning-bearing units* or *meaning units* were identified and coded according to their content [23]. Graneheim and Lundman [23] define *meaning units* as "words, sentences or paragraphs containing aspects related to each other through their content and context." Most comments were short and equaled one *meaning-bearing unit*; however, some contained more information and were represented by ≥ 2 *meaning-bearing units*. Therefore, the total number of coded *meaning-bearing units* was >186 .

The analysis began with each author performing a traditional deductive content analysis [23], independently categorizing the identified *meaning-bearing units* into categories. However, we quickly saw that the content of the comments went beyond traditional usability issues and matched the 8 dimensions of the sociotechnical framework by Sittig and Singh [24]. Therefore, we decided to use this framework to group the categories that emerged. The eight dimensions were (1) hardware and software computing infrastructure; (2) clinical content; (3) human-computer interface; (4) people; (5) workflow and communication; (6) internal organizational policies, procedures, and culture; (7) external rules, regulations, and pressures; and (8) system measurement and monitoring.

Ethics Approval

Data were collected from June to October 2016 after ethics approval of the research was granted by the Regional Ethical Review Board in Uppsala, Sweden (EPN 2016/129).

Results

Overview

During the survey period, 423,141 unique users logged in to Journalen, of whom 2587 (0.61%) patients completed the survey. Of all the respondents, 62.97% (1629/2587) identified as women and 30.85% (798/2587) as men; 0.39% (10/2587) of the respondents chose *other*, and 5.8% (150/2587) did not answer this question. According to use statistics provided by Inera AB, the company providing Journalen and the national patient portal [25], this reflects the gender distribution of the users in general (in 2016: 60% women and 40% men). Of all respondents, 39.81% (1030/2587) stated that they were working or had been

working within health care, and 54.54% (1411/2587) stated that they had no professional relation to health care; 5.64% (146/2587) of respondents did not answer this question. Respondents to this survey had a higher education level than the general population [17]. Among our respondents, 60.57% (1487/2455) had higher education compared with 42% of the general population [26]. Whether this is because users of Journalen are well educated or whether this is a subgroup of users who are more inclined to answer a survey is not known.

In summary, the survey results regarding user characteristics at the national level indicated that most respondents were women and that most had studied at least 3 years of higher education. In addition, the results indicate that many users of Journalen experienced both being patients and working as health care professionals.

In the overview of the survey results by Moll et al [17], details of the respondents' views on the usefulness and benefits of accessing their health records on the web are presented in more detail. Overall, the patients who answered the survey were positive toward Journalen. Respondents were asked to rate on a 5-point Likert scale the extent to which they agreed to the following more general statements: "I think that access to one's medical records online is generally a good reform" and "I think that access to Journalen is good for me."

Of all the respondents, $>96\%$ (2454/2541, 96.58%, and 2455/2528, 97.11%) for the respective questions showed a positive attitude toward Journalen (completely or partly agree). However, a positive attitude toward having access to one's health records does not say much about the usability of the system; therefore, we now present an analysis of the free-text comments related to the usability of the Swedish PAEHR.

Qualitative Analysis of Free-Text Answers

Overview

Analyzing free-text comments qualitatively adds to a deeper understanding of the usability issues patients experience when using a system, which, in this case, is Journalen. Of the 2587 respondents, 186 (7.19%) provided free-text comments on the usability questions. Most comments were short; however, some contained more information and were coded into several categories. A total of 9 comments did not include any useful information and were excluded from the presentation of the results.

The analysis resulted in 19 categories, which could be grouped under 7 of the 8 dimensions in the sociotechnical framework developed by Sittig and Singh [24] (Figure 1). Some of the categories had links to >1 dimension in the framework, as indicated by the lines in the figure. Although quantification of results is not common in qualitative analysis, we present the number of answers connected to each category in Table 1, as some categories were much more common in the findings than others.

The categories related to the sociotechnical domains are further described in the following sections.

Figure 1. Categories organized according to the sociotechnical framework.

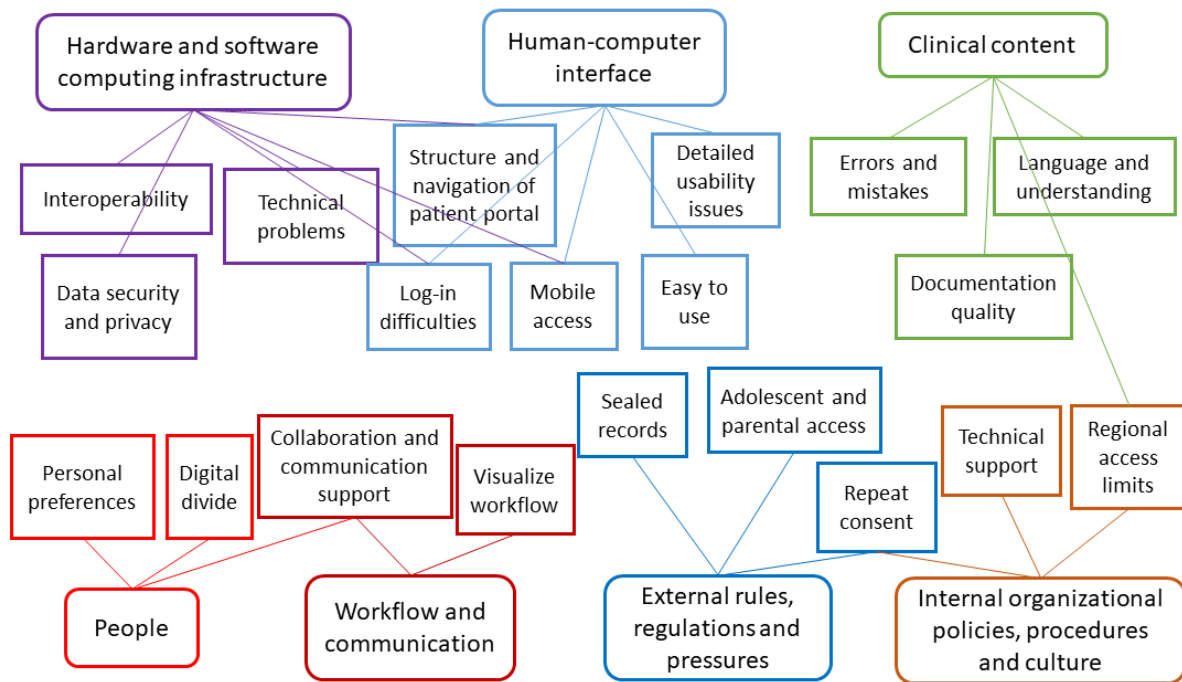


Table 1. Frequency of categories (N=186).

Sociotechnical domain and category	Frequency, n (%)
Hardware and software computing infrastructure	
Interoperability	3 (1.6)
Data security and privacy	2 (1.1)
Technical problems	9 (4.8)
Clinical content	
Errors and mistakes	4 (2.2)
Language and understanding	14 (7.5)
Documentation quality	3 (1.6)
Human-computer interface	
Structure and navigation of the patient portal	22 (11.8)
Log-in difficulties	6 (3.2)
Mobile access	5 (2.7)
Easy to use	3 (1.6)
Specific usability issues	7 (3.8)
People	
Personal preferences	3 (1.6)
Digital divide	6 (3.2)
Workflow and communication	
Collaboration and communication support	8 (4.3)
Visualize workflow	2 (1.1)
Internal organizational policies, procedures, and culture	
Regional access limitations	78 (41.9)
Technical support	2 (1.1)
External rules, regulations, and pressures	
Repeat consent	7 (3.8)
Adolescent and parental access	2 (1.1)
Sealed records	2 (1.1)

Hardware and Software Computing Infrastructure

Three categories were associated with the first domain: interoperability (3/186, 1.6%), data security and privacy (2/186, 1.1%), and technical problems (9/186, 4.8%). Interoperability issues raised by the respondents were mainly related to health care professionals not having access to information from other organizations, and the respondents expressed concern that this was the case:

Synchronize so that different regions can see each other's records. Now I have to deliver paper copies from one region to the other. Horrible, now that everything is digital.

A few respondents were concerned about how their data were protected and who had access to them:

I think there should be easy to find (short) information describing how my information is protected. It's funny that my entire record is available online without me

even knowing it...is it uploaded when I log in or is it there all the time?

Most comments in this domain were related to technical issues experienced by the respondents. Most described that the system was slow and appeared immature. Several comments also specifically highlighted access problems in a specific region.

Clinical Content

Three categories were associated with the clinical content: errors and mistakes (4/186, 2.2%), language and understanding (14/186, 7.5%), and documentation quality (3/186, 1.6%). Errors and mistakes included both positive and negative comments: most agreed that it is good that patients can find errors or mistakes, although some wished for easier ways of correcting such errors:

It is good to check the record from time to time, since there may be errors in the record that I can sort out together with healthcare.

Limited, it is not possible to write comments or change errors.

Approximately 7.5% (14/186) of comments in this category were related to medical language and understanding. Most of the comments related to difficulties in understanding the medical language and suggestions to include explanations of medical terms to facilitate users who do not have a health care background. However, there were also comments that highlighted that a patient can often understand the notes based on the context:

I'm surprised that it's so easy to understand what the doctors write. At least in my records, I've been easily able to follow what has happened and who did what. It increases my trust in healthcare.

Health care professionals made several of the comments, stating that they have no problem understanding the record themselves but that they worry that other patients may find it difficult. Finally, a few respondents complained that health care professionals did not use the correct structure or keywords when documenting or that there were discrepancies between documentation from different medical specialties. This could be considered a type of error or mistake but is expressed more as a form of poor documentation practice.

Human-Computer Interaction

Not surprisingly, the human-computer interaction dimension had 5 categories related to it: structure and navigation of the patient portal (22/186, 11.8%), log-in difficulties (6/186, 3.2%), mobile access (5/186, 2.7%), ease of use (3/186, 1.6%), and specific usability issues (7/186, 3.8%).

The structure and navigation of the patient portal was by far the largest category related to this dimension, and the issues described here concerned both general navigation issues (eg, too many clicks to reach Journalen) and challenges in understanding the relationship between different e-services offered on the patient portal:

Once you're in Journalen, its ok, but you have to understand the underlying structure to find the way...should I pick Journalen or Healthcare Events? Why can't I access the children's records from their account in the patient portal?

Some respondents also described challenges in logging into the patient portal. It is not necessary to sign up for portal access; however, one must have an electronic ID (eID) downloaded on their device. Most people in Sweden use a service provided by their bank; therefore, an eID is often referred to as a BankID. The eID can then be used to access several different public or private e-services, for example, to do tax returns, apply for parental or sick leave, or access web-based banking. However, understanding how to first download and install an eID can be challenging:

I would not have been able to get a BankID without help, or to access the patient portal without help. I have told a lot of people about this service, but they can't access it because they don't have a BankID.

Several comments were also related to the mobile interface. An increasing number of users access the patient portal using a mobile device (most often a smartphone), and there were complaints that the interface was poorly adapted for mobile access:

I think the user interface is a bit cumbersome and it is not always clear where you can find the information you're looking for. I usually use my mobile phone and the user interface is poorly adapted for mobile use. Perhaps a special mobile app would be good.

Although most respondents took the opportunity to describe aspects that needed improvement, a few also expressed that the system was easy to use:

Journalen seems more coherent and stable than the 8 healthcare IT-systems I have worked with professionally.

Finally, some respondents gave very specific feedback on usability issues they had encountered, for example, a link that did not work and data that needed to be entered in a specific format. An issue that was raised that could be especially important related to the function of allowing patients to request a prescription renewal. At the time, there was no connection between the form the patient filled in and the medication list; therefore, the patient had to manually enter the information about the prescription, increasing the risk for errors:

It would have been easier if one could choose an old prescription that needed renewal than to type everything again.

People

Some of the comments related to personal preferences (3/186, 1.6%) or the barriers to using digital services that exist, and how this type of eHealth service may not be used by many, potentially increasing the digital divide (6/186, 3.2%):

Waste of time, I'd rather discuss directly with my physician.

I'd prefer a paper record instead.

Most of the comments relating to the digital divide naturally voiced concern on behalf of others, as the respondents were obviously able to access Journalen themselves. Comments were related to socioeconomic factors, functional variations, age, and experience of using computers and digital tools:

The people I know that use the service are all middle class—no one from working class. I see a clear difference there.

The service is not accessible to people with certain types of disabilities and can be difficult for them to use.

Older people and those with limited experience of using computers will have a hard time to use this function, for some people it will be impossible. You should organize special information meetings with instructions for patients and their families.

Workflow and Communication

The dimension of workflow and communication had 2 categories: collaboration and communication support (8/186, 4.3%) and visualizing workflow (2/186, 1.1%).

Several respondents expressed a wish to use Journalen in a more collaborative way with their health care professionals, for example, by discussing the content in the record:

I wish the doctor took the initiative and discussed the information in the record with me. When I'm in the office with the physician my mind goes blank and it's difficult to talk about everything you'd like to, you realize after, and it feels like the doctor only gets a fragmented view of who you are. To discuss or add to the record could perhaps give us both a more complete picture.

Several respondents also wanted better tools for communicating directly with health care providers (eg, a chat function or secure emailing).

In one of the regions (Uppsala), users have access to a *log function*, which is a function allowing them to see whoever has accessed their record. Some users commented on how they used this function to see what was happening in their health care process and wanted even more information to keep themselves up to date with what was happening:

In the log, I would like to see the healthcare organization of the person who has accessed my record. Then I might be able to determine whether anything is happening with the referral that was sent, or if someone is checking up on what has happened to me as a patient since I was examined at their clinic.

Internal Organizational Policies, Procedures, and Culture

This dimension had only 2 categories, although one of them had by far the most comments: regional access limitations (78/186, 41.9%) and technical support (2/186, 1.1%). The latter had both a complaint and positive feedback.

Issues around regional access limitations mainly reflected that information users were expecting to find was missing. This is, of course, connected to the clinical content dimension as well; however, the reason some regions show more information than others is the regional adaptations of the national regulatory framework for PAEHR.

Most comments were related to certain types of information (most often laboratory results) that were missing from the record:

I want access to lab results, if they've taken a blood test e.g. — now I have to ask each doctor to print the results. I've brought these printouts to the primary care centre to facilitate communication.

A lot of functions are not available to me, and then they are "in the way" and a cause for disappointment. E.g., I don't have access to my lab results, which I would have appreciated, but the function is there but I can't access them. The system could have been

simpler and easier if only functions that are active are shown.

When information was perceived as missing in the record, some respondents also expressed concern that the quality of the documentation was poor and that the lack of documentation reduced their trust in health care:

A problem I experience is that healthcare doesn't seem to document properly in the record. I have a feeling that there is a real record...but I can only see part of it. Or the documentation is so poor that what I see is all there is. If that's the case, it worries me.

Another common complaint was that information was not retrospectively available. Many regions decided to only make data from the records available on the web from the date they went live with patients' access, whereas a few regions provided access to data retrospectively:

I thought more than the last month's data would be in the record...

Most regions also applied restrictions to certain clinical areas (eg, notes made in psychiatric care). A few respondents commented on this specifically, expressing that the lack of information made access less useful and that the specific blocking of mental health records felt discriminating:

I would use Journalen more if the whole record was included. The fact that the psychiatric notes are not included makes me feel discriminated and fragmented as a person. Body and mind affect each other and somatic care needs to consider what happens in psychiatry and vice versa.

Overall, the regional adaptations of the national regulatory framework caused both confusion and frustration, and when information was blocked, the respondents expressed that it reduced the usefulness and usability of the PAEHR.

External Rules, Regulations, and Pressures

Finally, a few categories were connected to external rules, regulations, and pressures: sealed records (2/186, 1.1%), adolescent and parental access (2/186, 1.1%), and repeat consent (7/186, 3.8%).

As a patient, one can request to have the whole or parts of their record to be sealed for access, hindering themselves from accessing it on the web or stopping the record from being accessible to other health care professionals involved in their care. This function was rarely used, although a few respondents who had sealed parts of their records expressed concern:

It's confusing when it says I can't access my record when I just wanted parts of it to be blocked for other healthcare providers.

The most frequent complaint in this category was related to having to consent to access the record every time the system was accessed. This was a great source of frustration for users; however, the design of the system took strict privacy and web-based access guidelines into account when designing the system:

I have to check a box that I have understood the same information every time I log in, the first time surely would have been enough?!

A few respondents commented on their access to their children's records. According to national regulations, legal guardians (most often parents) have automatic access to their children's records until they are aged 13 years. The age limit has been set in dialog with youth health care specialists and teenagers to protect adolescents' right to privacy. At the time of the survey, personal access to one's own record was granted only when they were aged 18 years:

It's bad that you can't see the children's records when they turn 13, and they can't log in themselves until they turn 18! A gap of 5 years without access to the online record!

Since then, the age for personal access has been lowered to 16 years; however, there is still a gap of 3 years between the loss of parental access and the activation of personal access.

Discussion

Principal Findings

To summarize, the results indicate that the respondents of the survey, who were also end users of the Swedish PAEHR Journalen, rated it fairly high on the SUS; however, many sociotechnical usability issues were identified through the respondents' free-text answers. Before discussing the results in more detail, we would like to address some of the methodological limitations of this study.

The Web-Based Survey

The survey distribution may have created a selection bias in the study, which should be considered when interpreting the results. The survey was distributed through the national patient portal and was only accessible to people who logged in and accessed the PAEHR. This was intentional, as the main aim of the study was to explore the experiences of people who had used the e-service; however, in doing so, we also excluded anyone who had previously struggled with poor usability and chosen not to continue using the service. Thus, users who were most critical to the service or its usability were likely not represented. If we had recruited individuals to represent the entire Swedish population, the results may have been different. Another potential bias was that the service at that time was not fully implemented throughout Sweden, and usability flaws that may occur for users who are patients in various regions may not have been detected. Thus, it is likely that the results would have covered other usability issues if the studied service would have been implemented simultaneously in all regions.

In addition, it was not possible to determine whether the participants of the survey were representative of all users of Journalen. As in most survey studies, the participants formed a small sample of all possible users, and many more users than those who answered the survey logged into Journalen during the 5 months when the survey was open. We do not know whether the survey respondents' demographic distribution is representative of all users of Journalen, and the subset of respondents included in this qualitative analysis was even

smaller. Despite this, we can assume that other users likely experienced the usability issues identified by our respondents as well.

Formal usability evaluations should be performed to complement these results to provide more details on the specific usability flaws encountered by end users and how these could best be addressed. However, capturing the daily frustrations of users related to, for example, limited access to certain types of information can be difficult in a traditional usability laboratory test where real patient data most likely cannot be used. Therefore, we argue that the results presented herein are of great importance.

Information Access Through a National Solution

The most frequent frustration commented on by the respondents was lack of access to specific information. We related this problem to local regulations, as each health care region is autonomous and thus has the opportunity to choose which information to share [27]. This proved to be not only difficult for end users to understand but also frustrating, potentially causing distrust in health care, as important information seemed to be missing from the record. Since then, a new regulatory framework has been agreed upon in Sweden, which states that all patients should have immediate access to all clinical documentation from tax-funded health care [28]. However, this change in the framework has not been enforced, and we continue to see differences between regions. Clinical notes from psychiatric care have, for example, often been excluded from patient access, not only in Sweden [29,30]. As our results indicated, excluding specific parts of the documentation can increase the sense of stigmatization and distrust. However, over the years, progress has been made, and when it comes to, for example, notes from psychiatric care, currently, 17 of 21 Swedish regions provide patients access [22] compared with only 2 when the survey was conducted [21].

PAEHRs as a Tool for Collaboration

An important aspect raised by some respondents was the need for more collaborative functions in the PAEHR. Using record access as a means of initiating and supporting dialog with their health care team was a wish expressed by several respondents. However, it appeared that the record was not used in this way by health care professionals [17]. When asked specifically about whether they as patients discussed the record content with health care professionals, 30.79% (766/2488) of the respondents completely or partly agreed, whereas 51.21% (1274/2488) did not agree or not at all agree. There still appears to be an underused opportunity to use the PAEHR for collaboration [31,32], especially considering that from the patient perspective, improved communication with health care professionals was stated as one of the most important reasons for accessing the PAEHR [17].

PAEHRs and the Digital Divide

Although we only reached actual users of the PAEHR Journalen through this survey, some comments were related to difficulties in logging in and using an eID. Some participants stated that they would not have been able to access the PAEHR without help, and others reported that friends and relatives were unable

to access the PAEHR as they did not have an eID. Low digital literacy is a well-known barrier to accessing and using eHealth [33] and is often correlated with older age or low education and socioeconomic status. Although we cannot draw any conclusions regarding this from our survey, the respondents in this study also had a higher education level than the general population. Although older users may experience more difficulties accessing and using the PAEHR, they also reported the greatest benefit of doing so [34]. Today, there is still very limited education or introduction of the PAEHR Journalen to patients; rather, they are expected to be able to use and understand the system intuitively. In fact, when asked whether they had been informed by health care professionals about the existence of Journalen, only 13.49% (335/2483) of the respondents agreed, and only 8.19% (203/2480) had been encouraged by health care professionals to read their records [17]. Although this may work well for many, making further educational resources available to patients may be a way of improving patients' experiences of using Journalen and could alleviate some of the problems described by the respondents in this study.

Since the survey was launched in 2016, the use of the Swedish PAEHR Journalen has increased rapidly, going from 77,000 individual users per month in January 2016 to 1,850,000 in January 2022 [25]. Uptake has been slow but steady, and the COVID-19 pandemic boosted use during 2020 and 2021 when, for example, COVID-19 test results were accessible through the PAEHR in most regions. We can assume that more people currently use both eIDs and the national patient portal, and a repeat survey is planned to take place in 2022, where we expect to reach a more diverse group of respondents.

PAEHRs and Proxy Access

Several respondents brought up proxy access, focusing specifically on parents' access to their children's records. The current regulation in Sweden automatically provides legal guardians (most often parents) access to a child's record until they are aged 13 years. The teenagers themselves can log in and access their records at the age of 16 years, leaving a 3-year gap when neither the child nor their parents have access to the record. This was critiqued by several respondents, as was the difficulty in finding the child's record in the patient portal. So far, limited research into the specific area of parental proxy and adolescent access has been performed in Sweden, and we know that regulations in this area differ between countries [2].

When it came to proxy access for other patient groups, there was a function for sharing parts of or the entire record with another person of the patient's own choosing at the time of the survey. However, this function was not widely used by the survey respondents [35], potentially as they managed their own health themselves. In 2018, the functionality was removed from the PAEHR as it was considered noncomplying with current legislation on the handling of patient data. A revision of the

legislation is currently underway, and functionality may be reopened in 2023. Further research into proxy access (both parental and other) is needed.

PAEHRs as a Tool for Patient Safety

Patient safety is often discussed in relation to patient portals and record access, with varying opinions on whether the effect is positive or negative. When patients are given web-based access to their records, it is not uncommon for them to discover errors or mistakes [36,37]. This was also commented upon by the respondents, where some used their access to ensure that no misunderstandings had occurred. However, the means for correcting errors were missing from the PAEHR, which was considered a problem by some respondents. Interesting research in the United States, where patients are provided with feedback tools to highlight and potentially correct errors, indicates that there could be real patient safety benefits from implementing such tools [38], something to be considered in other contexts as well.

Misunderstandings and confusion from accessing the records have been a concern for many who worry that patients may come to harm [39,40]. In this study, several respondents stated that the record could be difficult to understand and suggested, for example, that links to terminologies or explanations could be automatically added to the text. However, others also expressed that the records were surprisingly easy to understand. There can be several reasons why patients' experiences differ here: different levels of health literacy may affect their understanding, different medical specialties may use more or less difficult terminology, and individual health care professionals may express themselves in more or less easy-to-follow notes. Regardless, there appears to be room for improvement in this area, and both patient and health care professional education may be needed. How patient access may affect the way health care professionals actually document information is also an underexplored area [41], especially in domains such as mental health, where health care professionals have reported leaving important information out of the record as patients gain access to it [42,43].

Conclusions

Although the respondents of the survey regarding the PAEHR Journalen were overall satisfied with its usability [21], they also experienced important challenges when accessing their records. For all patients to be able to reap the benefits of record access, it is essential to understand both the usability challenges they encounter and, more broadly, how policies, regulations, and technical implementation decisions affect the usefulness of record access. The results presented here are specific to the Swedish PAEHR Journalen but also provide important insights into how the design and implementation of record access can be improved in any context.

Acknowledgments

This research was partly funded through the patient-centered assessment of patients' online access to electronic health records (PACESS) project (2016-00623) supported by FORTE—the Swedish Research Council for Health, Working Life and Welfare

- and partly supported by NordForsk through funding to Nordic eHealth for Patients: Benchmarking and Developing for the Future (project 100477). PACESS is hosted by the DOME consortium, which studies the Development of Online Medical Records and eHealth Services. We would also like to thank Inera AB for providing the data on the use of the national e-services and managing the survey and data collection through Journalen.

Authors' Contributions

Both authors were involved in the analysis and writing of the paper, and MH took the main responsibility. IS also participated in the survey design.

Conflicts of Interest

None declared.

References

1. Hägglund M, DesRoches C, Petersen C, Scandurra I. Patients' access to health records. *BMJ* 2019 Oct 02;367:15725. [doi: [10.1136/bmj.15725](https://doi.org/10.1136/bmj.15725)] [Medline: [31578180](https://pubmed.ncbi.nlm.nih.gov/31578180/)]
2. Essén A, Scandurra I, Gerrits R, Humphrey G, Johansen MA, Kierkegaard P, et al. Patient access to electronic health records: differences across ten countries. *Health Policy Technol* 2018 Mar;7(1):44-56. [doi: [10.1016/j.hlpt.2017.11.003](https://doi.org/10.1016/j.hlpt.2017.11.003)]
3. Cijvat CD, Cornet R, Hägglund M. Factors influencing development and implementation of patients' access to electronic health records—a comparative study of Sweden and the Netherlands. *Front Public Health* 2021 Jun 8;9:621210 [FREE Full text] [doi: [10.3389/fpubh.2021.621210](https://doi.org/10.3389/fpubh.2021.621210)] [Medline: [34169054](https://pubmed.ncbi.nlm.nih.gov/34169054/)]
4. Lyttkens L. Support Users To Access INformation and Services - Intermediate Pilot Evaluation Uppsala Version 1.1. Sustains. 2014 Jan 14. URL: <https://projectdome.files.wordpress.com/2021/01/sustains-rapport.pdf> [accessed 2020-10-02]
5. Tiik M. Rules and access rights of the Estonian integrated e-Health system. *Stud Health Technol Inform* 2010;156:245-256. [Medline: [20543358](https://pubmed.ncbi.nlm.nih.gov/20543358/)]
6. Wiljer D, Urowitz S, Apatu E, DeLenardo C, Eysenbach G, Harth T, Canadian Committee for Patient Accessible Health Records. Patient accessible electronic health records: exploring recommendations for successful implementation strategies. *J Med Internet Res* 2008 Oct 31;10(4):e34 [FREE Full text] [doi: [10.2196/jmir.1061](https://doi.org/10.2196/jmir.1061)] [Medline: [18974036](https://pubmed.ncbi.nlm.nih.gov/18974036/)]
7. Mandl KD, Szolovits P, Kohane IS. Public standards and patients' control: how to keep electronic medical records accessible but private. *BMJ* 2001 Feb 03;322(7281):283-287 [FREE Full text] [doi: [10.1136/bmj.322.7281.283](https://doi.org/10.1136/bmj.322.7281.283)] [Medline: [11157533](https://pubmed.ncbi.nlm.nih.gov/11157533/)]
8. Erlingsdottir G, Lindholm C. When patient empowerment encounters professional autonomy: the conflict and negotiation process of inscribing an eHealth service. *Scand J Public Adm* 2015;19(2):27-48. [doi: [10.1007/978-1-137-54604-3_5](https://doi.org/10.1007/978-1-137-54604-3_5)]
9. Delbanco T, Walker J, Darer JD, Elmore JG, Feldman HJ, Leveille SG, et al. Open notes: doctors and patients signing on. *Ann Intern Med* 2010 Jul 20;153(2):121-125 [FREE Full text] [doi: [10.7326/0003-4819-153-2-201007200-00008](https://doi.org/10.7326/0003-4819-153-2-201007200-00008)] [Medline: [20643992](https://pubmed.ncbi.nlm.nih.gov/20643992/)]
10. Delbanco T, Walker J, Bell SK, Darer JD, Elmore JG, Farag N, et al. Inviting patients to read their doctors' notes: a quasi-experimental study and a look ahead. *Ann Intern Med* 2012 Oct 02;157(7):461-470 [FREE Full text] [doi: [10.7326/0003-4819-157-7-201210020-00002](https://doi.org/10.7326/0003-4819-157-7-201210020-00002)] [Medline: [23027317](https://pubmed.ncbi.nlm.nih.gov/23027317/)]
11. Walker J, Leveille S, Bell S, Chimowitz H, Dong Z, Elmore JG, et al. OpenNotes after 7 years: patient experiences with ongoing access to their clinicians' outpatient visit notes. *J Med Internet Res* 2019 May 06;21(5):e13876 [FREE Full text] [doi: [10.2196/13876](https://doi.org/10.2196/13876)] [Medline: [31066717](https://pubmed.ncbi.nlm.nih.gov/31066717/)]
12. Salmi L, Blease C, Hägglund M, Walker J, DesRoches CM. US policy requires immediate release of records to patients. *BMJ* 2021 Feb 18;372:n426. [doi: [10.1136/bmj.n426](https://doi.org/10.1136/bmj.n426)] [Medline: [33602667](https://pubmed.ncbi.nlm.nih.gov/33602667/)]
13. 21st Century Cures Act: Interoperability, Information Blocking, and the ONC Health IT Certification Program. Federal Register – The Daily Journal of United States Government. 2020 Jan 5. URL: <https://www.federalregister.gov/documents/2020/05/01/2020-07419/21st-century-cures-act-interoperability-information-blocking-and-the-onc-health-it-certification> [accessed 2020-12-30]
14. Hägglund M, Scandurra I. Patients' online access to electronic health records: current status and experiences from the implementation in Sweden. *Stud Health Technol Inform* 2017;245:723-727. [Medline: [29295193](https://pubmed.ncbi.nlm.nih.gov/29295193/)]
15. Davoody N, Koch S, Krakau I, Hägglund M. Accessing and sharing health information for post-discharge stroke care through a national health information exchange platform - a case study. *BMC Med Inform Decis Mak* 2019 May 03;19(1):95 [FREE Full text] [doi: [10.1186/s12911-019-0816-x](https://doi.org/10.1186/s12911-019-0816-x)] [Medline: [31053141](https://pubmed.ncbi.nlm.nih.gov/31053141/)]
16. Sellberg N, Eltes J. The Swedish patient portal and its relation to the national reference architecture and the overall eHealth infrastructure. In: Aanstad M, Grisot M, Hanseth O, Vassilakopoulou P, editors. *Information Infrastructures within European Health Care: Working with the Installed Base*. Cham, Switzerland: Springer; 2017:225-244.
17. Moll J, Rexhepi H, Cajander Å, Grünloh C, Huvila I, Hägglund M, et al. Patients' experiences of accessing their electronic health records: national patient survey in Sweden. *J Med Internet Res* 2018 Nov 01;20(11):e278 [FREE Full text] [doi: [10.2196/jmir.9492](https://doi.org/10.2196/jmir.9492)] [Medline: [30389647](https://pubmed.ncbi.nlm.nih.gov/30389647/)]

18. Abd-Alrazaq AA, Bewick BM, Farragher T, Gardner P. Factors that affect the use of electronic personal health records among patients: a systematic review. *Int J Med Inform* 2019 Jun;126:164-175. [doi: [10.1016/j.ijmedinf.2019.03.014](https://doi.org/10.1016/j.ijmedinf.2019.03.014)] [Medline: [31029258](https://pubmed.ncbi.nlm.nih.gov/31029258/)]
19. Bangor A, Kortum PT, Miller JT. An empirical evaluation of the system usability scale. *Int J Hum Comput Interact* 2008 Jul 30;24(6):574-594. [doi: [10.1080/10447310802205776](https://doi.org/10.1080/10447310802205776)]
20. Brooke J. SUS: a 'quick and dirty' usability scale. In: Jordan PW, Thomas B, McClelland IL, Weerdmeester B, editors. *Usability Evaluation in Industry*. London, UK: Taylor & Francis; 1996:189-194.
21. Hägglund M, Scandurra I. User evaluation of the Swedish patient accessible electronic health record: system usability scale. *JMIR Hum Factors* 2021 Jul 27;8(3):e24927 [FREE Full text] [doi: [10.2196/24927](https://doi.org/10.2196/24927)] [Medline: [34313596](https://pubmed.ncbi.nlm.nih.gov/34313596/)]
22. Bärkås A, Scandurra I, Rexhepi H, Blease C, Cajander Å, Hägglund M. Patients' access to their psychiatric notes: current policies and practices in Sweden. *Int J Environ Res Public Health* 2021 Aug 30;18(17):9140 [FREE Full text] [doi: [10.3390/ijerph18179140](https://doi.org/10.3390/ijerph18179140)] [Medline: [34501730](https://pubmed.ncbi.nlm.nih.gov/34501730/)]
23. Graneheim UH, Lundman B. Qualitative content analysis in nursing research: concepts, procedures and measures to achieve trustworthiness. *Nurse Educ Today* 2004 Feb;24(2):105-112. [doi: [10.1016/j.nedt.2003.10.001](https://doi.org/10.1016/j.nedt.2003.10.001)] [Medline: [14769454](https://pubmed.ncbi.nlm.nih.gov/14769454/)]
24. Sittig DF, Singh H. A new sociotechnical model for studying health information technology in complex adaptive healthcare systems. *Qual Saf Health Care* 2010 Oct;19 Suppl 3:i68-i74 [FREE Full text] [doi: [10.1136/qshc.2010.042085](https://doi.org/10.1136/qshc.2010.042085)] [Medline: [20959322](https://pubmed.ncbi.nlm.nih.gov/20959322/)]
25. Statistik för Ineras tjänster. Inera. URL: <https://www.inera.se/aktuellt/statistik/> [accessed 2021-11-30]
26. Educational attainment of the population in 2016: the Stockholm area has the highest level of education. Statistics Sweden. 2017 Jun 21. URL: <https://www.scb.se/en/finding-statistics/statistics-by-subject-area/education-and-research/education-of-the-population/educational-attainment-of-the-population/pong/statistical-news/educational-attainment-of-the-population-in-2016/> [accessed 2021-07-15]
27. Scandurra I, Lyttkens L, Eklund B. Implications of Swedish national regulatory framework of the patient accessible electronic health record. *Stud Health Technol Inform* 2016;228:695-699. [Medline: [27577474](https://pubmed.ncbi.nlm.nih.gov/27577474/)]
28. Scandurra I, Pettersson M, Eklund B, Lyttkens L. Analysis of the updated Swedish regulatory framework of the patient accessible electronic health record in relation to usage experience. *Stud Health Technol Inform* 2017;245:798-802. [Medline: [29295208](https://pubmed.ncbi.nlm.nih.gov/29295208/)]
29. Blease CR, O'Neill S, Walker J, Hägglund M, Torous J. Sharing notes with mental health patients: balancing risks with respect. *Lancet Psychiatry* 2020 Nov;7(11):924-925. [doi: [10.1016/S2215-0366\(20\)30032-8](https://doi.org/10.1016/S2215-0366(20)30032-8)] [Medline: [32059796](https://pubmed.ncbi.nlm.nih.gov/32059796/)]
30. Schwarz J, Bärkås A, Blease C, Collins L, Hägglund M, Markham S, et al. Sharing clinical notes and electronic health records with people affected by mental health conditions: scoping review. *JMIR Ment Health* 2021 Dec 14;8(12):e34170 [FREE Full text] [doi: [10.2196/34170](https://doi.org/10.2196/34170)] [Medline: [34904956](https://pubmed.ncbi.nlm.nih.gov/34904956/)]
31. Ålander T, Scandurra I. Experiences of healthcare professionals to the introduction in Sweden of a public eHealth service: patients' online access to their electronic health records. *Stud Health Technol Inform* 2015;216:153-157. [Medline: [26262029](https://pubmed.ncbi.nlm.nih.gov/26262029/)]
32. Scandurra I, Jansson A, Forsberg-Fransson ML, Ålander T. Patient accessible EHR is controversial: lack of knowledge and diverse perceptions among professions. *Int J Reliab Qual E-Healthcare* 2017 Jan;6(1):29-45 [FREE Full text] [doi: [10.4018/ijrqeh.2017010103](https://doi.org/10.4018/ijrqeh.2017010103)]
33. Hong YA, Cho J. Has the digital health divide widened? Trends of health-related internet use among older adults from 2003 to 2011. *J Gerontol B Psychol Sci Soc Sci* 2017 Sep 01;72(5):856-863. [doi: [10.1093/geronb/gbw100](https://doi.org/10.1093/geronb/gbw100)] [Medline: [27558403](https://pubmed.ncbi.nlm.nih.gov/27558403/)]
34. Huvila I, Cajander Å, Moll J, Enwald H, Eriksson-Backa K, Rexhepi H. Technological and informational frames: explaining age-related variation in the use of patient accessible electronic health records as technology and information. *Inf Technol People* 2022;35(8):1-22. [doi: [10.1108/itp-08-2020-0566](https://doi.org/10.1108/itp-08-2020-0566)]
35. Nurgalieva L, Cajander Å, Moll J, Åhlfeldt RM, Huvila I, Marchese M. 'I do not share it with others. No, it's for me, it's my care': on sharing of patient accessible electronic health records. *Health Informatics J* 2020 Dec;26(4):2554-2567 [FREE Full text] [doi: [10.1177/1460458220912559](https://doi.org/10.1177/1460458220912559)] [Medline: [32264734](https://pubmed.ncbi.nlm.nih.gov/32264734/)]
36. Bell SK, Delbanco T, Elmore JG, Fitzgerald PS, Fossa A, Harcourt K, et al. Frequency and types of patient-reported errors in electronic health record ambulatory care notes. *JAMA Netw Open* 2020 Jun 01;3(6):e205867 [FREE Full text] [doi: [10.1001/jamanetworkopen.2020.5867](https://doi.org/10.1001/jamanetworkopen.2020.5867)] [Medline: [32515797](https://pubmed.ncbi.nlm.nih.gov/32515797/)]
37. Bell SK, Mejilla R, Anselmo M, Darer JD, Elmore JG, Leveille S, et al. When doctors share visit notes with patients: a study of patient and doctor perceptions of documentation errors, safety opportunities and the patient-doctor relationship. *BMJ Qual Saf* 2017 Apr;26(4):262-270 [FREE Full text] [doi: [10.1136/bmjqs-2015-004697](https://doi.org/10.1136/bmjqs-2015-004697)] [Medline: [27193032](https://pubmed.ncbi.nlm.nih.gov/27193032/)]
38. Bell SK, Gerard M, Fossa A, Delbanco T, Folcarelli PH, Sands KE, et al. A patient feedback reporting tool for OpenNotes: implications for patient-clinician safety and quality partnerships. *BMJ Qual Saf* 2017 Apr;26(4):312-322. [doi: [10.1136/bmjqs-2016-006020](https://doi.org/10.1136/bmjqs-2016-006020)] [Medline: [27965416](https://pubmed.ncbi.nlm.nih.gov/27965416/)]
39. Grünloh C, Cajander Å, Myreteg G. "The record is our work tool!"-physicians' framing of a patient portal in Sweden. *J Med Internet Res* 2016 Jun 27;18(6):e167 [FREE Full text] [doi: [10.2196/jmir.5705](https://doi.org/10.2196/jmir.5705)] [Medline: [27349531](https://pubmed.ncbi.nlm.nih.gov/27349531/)]

40. de Lusignan S, Mold F, Sheikh A, Majeed A, Wyatt JC, Quinn T, et al. Patients' online access to their electronic health records and linked online services: a systematic interpretative review. *BMJ Open* 2014 Sep 08;4(9):e006021 [[FREE Full text](#)] [doi: [10.1136/bmjopen-2014-006021](https://doi.org/10.1136/bmjopen-2014-006021)] [Medline: [25200561](https://pubmed.ncbi.nlm.nih.gov/25200561/)]
41. Blease C, Torous J, Hägglund M. Does patient access to clinical notes change documentation? *Front Public Health* 2020 Nov 27;8:577896 [[FREE Full text](#)] [doi: [10.3389/fpubh.2020.577896](https://doi.org/10.3389/fpubh.2020.577896)] [Medline: [33330320](https://pubmed.ncbi.nlm.nih.gov/33330320/)]
42. Petersson L, Erlingsdóttir G. Open Notes in Swedish psychiatric care (part 2): survey among psychiatric care professionals. *JMIR Ment Health* 2018 Jun 21;5(2):e10521 [[FREE Full text](#)] [doi: [10.2196/10521](https://doi.org/10.2196/10521)] [Medline: [29929946](https://pubmed.ncbi.nlm.nih.gov/29929946/)]
43. Kristiansen E, Johansen M, Zanaboni P. Healthcare personnels' experience with patients' online access to electronic health records: Differences between professions, regions, and somatic and psychiatric healthcare. In: *Proceedings of the 17th Scandinavian Conference on Linköping Electronic Conference*. 2019 Presented at: SHI '19; November 12-13, 2019; Oslo, Norway p. 93-98.

Abbreviations

EHR: electronic health record

PAEHR: patient-accessible electronic health record

SUS: System Usability Scale

eID: electronic ID

Edited by A Kushniruk; submitted 13.02.22; peer-reviewed by J Schwarz, S Markham, R Marshall; comments to author 23.04.22; revised version received 06.05.22; accepted 14.05.22; published 23.06.22.

Please cite as:

Hägglund M, Scandurra I

Usability of the Swedish Accessible Electronic Health Record: Qualitative Survey Study

JMIR Hum Factors 2022;9(2):e37192

URL: <https://humanfactors.jmir.org/2022/2/e37192>

doi: [10.2196/37192](https://doi.org/10.2196/37192)

PMID: [35737444](https://pubmed.ncbi.nlm.nih.gov/35737444/)

©Maria Hägglund, Isabella Scandurra. Originally published in *JMIR Human Factors* (<https://humanfactors.jmir.org>), 23.06.2022. This is an open-access article distributed under the terms of the Creative Commons Attribution License (<https://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in *JMIR Human Factors*, is properly cited. The complete bibliographic information, a link to the original publication on <https://humanfactors.jmir.org>, as well as this copyright and license information must be included.

Original Paper

The Sociodemographic Digital Divide in Mobile Health App Use Among Clients at Outpatient Departments in Inner Mongolia, China: Cross-sectional Survey Study

Li Cao^{1,2}, MD; Virasakdi Chongsuvivatwong², MD, PhD; Edward B McNeil², MSc

¹Information Technology Department, Inner Mongolia Medical University, Hohhot, China

²Department of Epidemiology, Faculty of Medicine, Prince of Songkla University, Songkhla, Thailand

Corresponding Author:

Virasakdi Chongsuvivatwong, MD, PhD

Department of Epidemiology

Faculty of Medicine

Prince of Songkla University

Kanjanavanich Road, Hat Yai District

Songkhla, 90110

Thailand

Phone: 66 74451165

Email: cvirasak@medicine.psu.ac.th

Abstract

Background: Mobile health (mHealth) apps have become part of the infrastructure for access to health care in hospitals, especially during the COVID-19 pandemic. However, little is known about the effects of sociodemographic characteristics on the digital divide regarding the use of hospital-based mHealth apps and their benefits to patients and caregivers.

Objective: The aim of this study was to document the cascade of potential influences from digital access to digital use and then to mHealth use, as well as the potential influence of sociodemographic variables on elements of the cascade.

Methods: A cross-sectional survey was conducted from January to February 2021 among adult clients at outpatient departments in 12 tertiary hospitals of Inner Mongolia, China. Structural equation modeling was conducted after the construct comprising digital access, digital use, and mHealth use was validated.

Results: Of 2115 participants, the β coefficients (95% CI) of potential influence of digital access on digital use, and potential influence of digital use on mHealth use, were 0.28 (95% CI 0.22-0.34) and 0.51 (95% CI 0.38-0.64), respectively. Older adults were disadvantaged with regard to mHealth access and use ($\beta=-0.38$ and $\beta=-0.41$), as were less educated subgroups ($\beta=-0.24$ and $\beta=-0.27$), and these two factors had nonsignificant direct effects on mHealth use.

Conclusions: To overcome the mHealth use divide, it is important to improve digital access and digital use among older adults and less educated groups.

(*JMIR Hum Factors* 2022;9(2):e36962) doi:[10.2196/36962](https://doi.org/10.2196/36962)

KEYWORDS

digital divide; mHealth; app; structural equation modeling; client

Introduction

The term mobile health (mHealth) was coined in 2003 [1] and is defined as health care practice through mobile devices and their apps [2,3]. mHealth apps have been developed for hospitals to allocate and manage their medical care services and to improve patient satisfaction [4,5]. During the COVID-19 pandemic, mHealth apps were used to implement prescreening, tracking cases, and social distancing measures [6-8]. COVID-19

as an extra factor is also exacerbating existing inequalities [9]. Older adults and less educated people have been affected the most by lockdown measures [9,10]. Some hospitals have implemented mHealth-based online appointments and video consultations with health care providers [11,12], instead of traditional register windows and consultation rooms, in order to reduce contacts [13,14]. Vulnerable people benefit the least from these digital solutions [9,10].

The digital divide is defined as a gap between people who have access to internet services and those who do not [15,16]. The digital divide is a central issue in the world today [17,18] and is described using a three-level framework [19] (Figure 1).

Around half the number of people worldwide have access to the internet. Sociodemographic characteristics, particularly age, gender, education, and income, predict internet access and use [7,20].

Figure 1. The three-level digital divide framework.



As of December 2020, Chinese internet penetration had reached 70.4% [21]. Patient satisfaction has become an important indicator for measuring health care quality [22,23] and policy evaluations of health care systems, which directly connect with health services use [24]. mHealth app features affect user satisfaction in various health care scenarios [25,26]. mHealth is a specific area that can be used to examine the digital divide. The Chinese government has tried to implement reforms to reduce waiting times and improve health care and patient satisfaction, including the use of mHealth [27-29]. In 2015, the Chinese government issued an action plan requiring at least 50% of appointments to be online for visiting doctors in tertiary hospitals by the end of 2017 [30]. However, the adoption rate of mobile services for outpatients was low, accounting for only 31.5% in 2019 [31]. Whether mHealth in China has reached its goal is still a subject for debate.

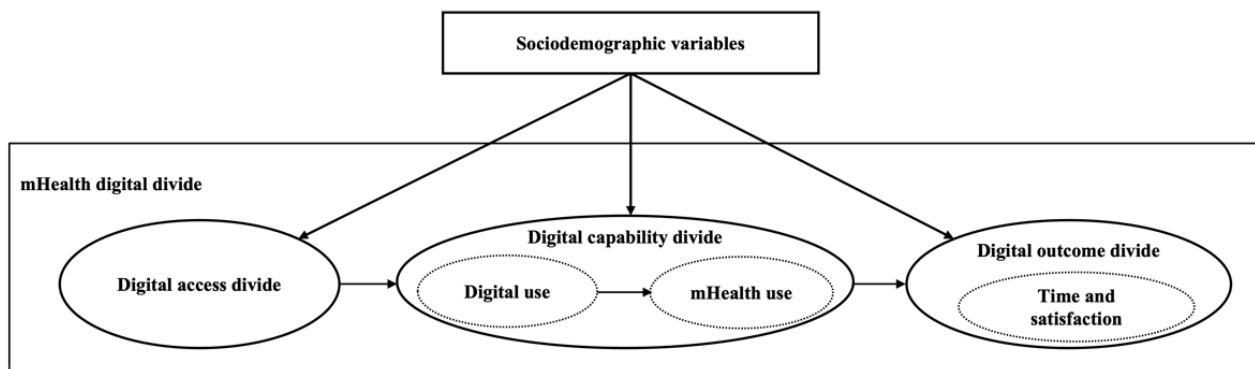
access to medical care services and health outcomes [35,36]. Meanwhile, hospital-based mHealth apps can be used by caregivers, who may act as proxy users on behalf of patients to reduce the digital divide [37].

In China, 83% of tertiary hospitals provide online appointments, of which 60% have mHealth services [4]. The majority of mHealth apps were nested into WeChat using an official account or a mini-app [12,32]; WeChat is the most popular social media platform in China, whereas other platforms have been built by local governments, companies, or hospitals. Common mHealth services based out of hospitals are extremely similar to each other. Generally, potential users must first install the mHealth app or subscribe to the WeChat official account, sign up to be a user, then sign in to use the mHealth service. Users can make an appointment with the doctors and pay the fee for clinical tests or medicine [33]. mHealth use is recognized as a fundamental social determinant of health [34], which facilitates

Inner Mongolia is a province located in the northern part of China and borders Mongolia and Russia to the north. It is an underdeveloped province [38] containing 49 ethnic minority groups. Traditional Mongolian medicine, traditional Chinese medicine, and Western medicine are well supported by the government and accepted by local citizens [39]. Within this context, it is likely that there is digital divide between different groups of health care users. A study on digital divide among hospital clients in this area could lead to improvement of health care in the future.

To date, few studies have been conducted to analyze the digital divide regarding the use of hospital-based mHealth apps. All previous studies ran regressions to find the potential influence of sociodemographic variables on mHealth use without considering the clients' digital access and digital use backgrounds. Our framework of evaluation of the digital divide in mHealth covers the whole spectrum, from digital access to digital use, mHealth use, and time and satisfaction when using health care services (Figure 2). Separating the potential influence of sociodemographics on each part of the cascade, with simultaneous evaluation of the flow of potential influence along the cascade, will lead to better understanding and a more appropriate formulation of policy to minimize these digital divides.

Figure 2. Research and hypothesis model of mobile health (mHealth) digital divide.



Methods

Study Design and Setting

A cross-sectional survey was conducted in January and February 2021 in 12 tertiary hospitals across three cities of the Inner Mongolia Autonomous Region. All studied hospitals provide mHealth apps with eHealth codes, appointments with doctors, electronic payment (e-payment), and health record checking.

Participants

Clients (ie, patients and caregivers) who were visiting the outpatient department for nonemergency care, were aged 18 years or older, and able to speak Chinese were eligible for the study.

Questionnaire

The questionnaire was created by the research team and reviewed by an epidemiologist from the public health school with mHealth research expertise. The IT department of the Affiliated People's Hospital of Inner Mongolia Medical University was consulted twice on the amended questionnaire.

Data Collection Procedure

A study team from the Inner Mongolia Medical University comprised of resident physicians was assembled and trained on the data collection process. The interviewers consecutively contacted clients at a drugstore or outpatient departure areas, explained the study, and asked for their consent to participate in the study. Consenting participants were asked to complete the questionnaire via face-to-face interviews.

Ethics Approval

This research study was approved by the Office of Human Research Ethics Committee, Faculty of Medicine, Prince of Songkla University, Thailand (REC.63-306-18-1).

Variables

Independent Variables

Independent variables included demographic information (ie, age, gender, and area of residence) and socioeconomic status (ie, educational level, employment status, and household monthly income).

Constructing the Mediators

Initially, we created a set of digital activities, including having household internet bandwidth, using a computer, using email, having a smartphone, having the ability to install apps, having wearable devices, using the internet, shopping online, using e-payment, using social media, searching online, using the internet daily, and using the internet for more than 5 years. For mHealth app use, since all apps were similar, we did not specify an mHealth app name in the questionnaire. mHealth app use was a latent intermediary variable, with its value loaded onto a series of observed variables: having mHealth apps, having an eHealth code, making appointments with doctors, attending consultations with doctors, making e-payments for medical care, and checking health records. All of these observed variables were on a "yes or no" binary scale. We then randomly split the

data set in half to run exploratory factor analysis (EFA) in order to obtain the constructs; we then analyzed reliability within each factor. Subsequently, the constructs were tested on the remaining half of the data set for validation via confirmatory factor analysis (CFA).

Dependent Variables

Dependent variables included time and satisfaction with health care. We selected four indicators related to mHealth use: waiting time, check-in process, medicine withdrawal and payment process, and general satisfaction. These indicators were assessed using the Chinese Outpatient Experience Questionnaire [40], which uses a 5-point Likert scale, ranging from 1 (the worst satisfaction) to 5 (the best satisfaction).

Statistical Analysis

Data were double-entered into EpiData (version 3.1; The EpiData Association), and analysis was performed using R (version 4.1.0; The R Foundation). Descriptive statistics were used to summarize the characteristics of the clients, namely frequency with percentage for categorical variables and mean with SD for continuous variables.

CFA was used to analyze the correlation matrix among the domains. A multiple-indicator, multiple-cause model (MIMIC) with structural equation modeling (SEM) [41] was used to examine the association between sociodemographic variables, mHealth use, and time and satisfaction with health care. The "polycor" R package was used for polychoric and polyserial correlations of categorical variables [42], the "psych" R package was used for EFA [43], and the "lavaan" R package was used for CFA and SEM [44]. The sample size per hospital was calculated based on the assumption that 38.4% of Chinese adults have an mHealth app [45], using an infinite population proportion formula as follows:

$$n = \frac{Z^2 \cdot p \cdot q}{d^2}$$

with a 10% error rate (d) and a 95% CI ($\alpha=.05$). A 10% nonresponse rate was also assumed. With these parameters, 102 participants were required to be recruited from each hospital. Due to the effect of COVID-19 on mHealth use, we decided to recruit 200 participants from each hospital. Finally, 2366 participants were included.

Results

Sociodemographic Factors

A total of 2115 clients provided valid responses. Their mean age was 43.34 (SD 15.39) years. Other demographic characteristics are summarized in Table 1. The participants were distributed nearly equally between the two genders. Three-quarters resided in an urban area. Almost half of the participants were educated at the tertiary education level. More than half were employed by the government or a company. Their household incomes were also somewhat evenly distributed, with a median of ¥4000 to ¥6000 (a currency exchange rate of ¥1=US \$0.15 is applicable), which was deemed to be middle class in China [46].

Table 1. Basic characteristics of the participants.

Variable	Participants (N=2115), n (%) ^a
Sociodemographic variables	
Age (years), mean (SD)	43.34 (15.39)
Gender	
Male	1007 (47.61)
Female	1108 (52.39)
Urban residence	
Yes	1630 (77.07)
No	485 (22.93)
Educational level	
Primary or less	297 (14.04)
Secondary	805 (38.06)
Tertiary	1013 (47.90)
Employment status	
Employed	1166 (55.13)
Unemployed	949 (44.87)
Household monthly income (¥^b)	
0-2000	220 (10.40)
2001-4000	424 (20.05)
4001-6000	456 (21.56)
6001-8000	345 (16.31)
8001-9999	314 (14.85)
≥10,000	356 (16.83)
Digital activities	
Have household bandwidth	
No	365 (17.26)
Yes	1750 (82.74)
Use a computer	
No	669 (31.63)
Yes	1446 (68.37)
Use email	
No	925 (43.74)
Yes	1190 (56.26)
Have a smartphone	
No	129 (6.10)
Yes	1986 (93.90)
Have the ability to install apps	
No	584 (27.61)
Yes	1531 (72.39)
Have wearable devices	
No	1564 (73.95)
Yes	551 (26.05)

Variable	Participants (N=2115), n (%) ^a
Use the internet	
No	244 (11.54)
Yes	1871 (88.46)
Shop online	
No	555 (26.24)
Yes	1560 (73.76)
Use e-payment^c	
No	429 (20.28)
Yes	1686 (79.72)
Use social media	
No	548 (25.91)
Yes	1567 (74.09)
Search online	
No	543 (25.67)
Yes	1572 (74.33)
Daily internet use	
No	473 (22.36)
Yes	1642 (77.64)
More than 5 years of internet use	
No	687 (32.48)
Yes	1428 (67.52)
mHealth^d use	
Have mHealth apps	
No	792 (37.45)
Yes	1323 (62.55)
Have eHealth code	
No	682 (32.25)
Yes	1433 (67.75)
Make appointments with doctors	
No	978 (46.24)
Yes	1137 (53.76)
Attend consultations with doctors	
No	1876 (88.70)
Yes	239 (11.30)
Use e-payment for medical care	
No	1102 (52.10)
Yes	1013 (47.90)
Health record checking	
No	1409 (66.62)

Variable	Participants (N=2115), n (%) ^a
Yes	706 (33.38)

^aAll values are reported as n (%), except for the age variable.

^bA currency exchange rate of ¥1=US \$0.15 is applicable.

^ce-payment: electronic payment.

^dmHealth: mobile health.

Digital Activities

Overall, 88.46% of the participants used the internet, 82.74% had access to the internet at home, and 93.90% had a smartphone. In total, 68.37% of participants used a computer, and 56.26% could use email. Three-quarters of the participants could self-install an app, and one-quarter wore smart wearable devices. Around three-quarters of the participants purchased commodities online, used e-payment, used social media, and performed information-searching online. Most used the internet daily and had been using it for more than 5 years.

mHealth App Use

Overall, 62.55% of the participants had an mHealth app, 67.75% had an eHealth code, 53.76% could make an appointment to see a doctor, 47.90% used e-payment for health care, 33.38% reviewed their health record on an mHealth app, and 11.30% consulted with a doctor online (Table 1).

EFA Model of mHealth Digital Divide

From the EFA, validation of the classification of digital activities was performed with a Kaiser-Meyer-Olkin test [47], with a

sample adequacy of 0.929 [48], and a Bartlett test of sphericity, which was statistically significant ($\chi^2_{253}=1260.1, P<.001$) [49]. Based on the parallel analysis, four factors were determined [50]. Due to the nonnormally distributed data, principal axis factoring was used as an appropriate extraction method [51], and oblimin rotation was used as an appropriate oblique rotation method [52]. Five items (ie, “have household bandwidth,” “have wearable devices,” “use social media,” “use e-payment,” and “attend consultations with doctors”) had factor loadings of less than 0.4 [53] or high cross-loadings [54] and were, thus, dropped from the analysis.

Finally, we came up with four domains (ie, factors): digital access, digital use, mHealth use, and time and satisfaction with health care. Details of the EFA are shown in Table 2.

The factor loadings were high, ranging from 0.469 to 0.886. Cronbach α values were all above .8. These four factors explained 62% of the variance. All of these values suggested that our construct was adequate.

Table 2. Measurement items and their reliability by exploratory factor analysis.

Factor and items	Loading	Communality	Cronbach α^a	Proportion of total variance ^a
Digital access				
Have a smartphone	0.717	0.557	.80	0.121
Use the internet	0.886	0.825		
Daily internet use	0.542	0.478		
Digital use				
Have the ability to install apps	0.740	0.716	.92	0.212
More than 5 years of internet use	0.613	0.517		
Shop online	0.604	0.676		
Search online	0.696	0.689		
Use a computer	0.878	0.721		
Use email	0.881	0.703		
mHealth^b use				
Have mHealth apps	0.610	0.533	.86	0.154
Have eHealth code	0.469	0.373		
Make appointments with doctors	0.863	0.646		
Use e-payment ^c for medical care	0.846	0.771		
Health record checking	0.710	0.563		
Time and satisfaction with health care				
Waiting time	0.779	0.586	.84	0.133
Check-in process	0.836	0.728		
Medicine withdrawal and payment process	0.803	0.658		
General satisfaction	0.649	0.423		

^aValues for groups are reported in the row of the top variable of the group.

^bmHealth: mobile health.

^ce-payment: electronic payment

CFA and SEM Model of mHealth Digital Divide

The second part of the data set was used to assess reliability and validity by CFA. The measurement model of digital divide in mHealth app use was adequately measured by associated indicators based on high factor loadings. The correlations among latent variables by CFA are shown in Table 3. The model fitted the data well, according to the following indices: $\chi^2_{129}=528.3$, $\chi^2/df=4.095$, comparative fit index (CFI)=0.940, Tucker-Lewis index (TLI)=0.929, root mean square error of approximation (RMSEA)=0.054 (90% CI 0.049-0.059), and standardized root mean square residual (SRMR)=0.042. The Cronbach α reliability coefficient was greater than .7, and convergent validity based on average variance extracted was greater than 0.5.

A MIMIC model with SEM was investigated for the model of mHealth digital divide by using the weighted least square mean and variance adjusted estimator [55], since most of the variables were categorical. The overall indices of the final SEM model fitted the data well: $\chi^2_{216}=682.6$, $\chi^2/df=3.160$, RMSEA=0.045

(90% CI 0.041-0.049), SRMR=0.036, CFI=0.949, and TLI=0.938 [53].

Figure 3 shows the regressions of all the paths. The details of the β coefficients and 95% CIs are shown in Table 4. Age was taken as a continuous variable. Although education and household income were initially ordinal categorical variables, we standardized them as continuous variables to suit the SEM. Their coefficients were interpreted as whether these had a dose-response relationship with the outcome. Age and education were strongly associated with digital access and digital use. Income had a low effect on digital access, and income, employment status, and urban residence were weakly correlated with digital use. No significant gender gap regarding these variables was seen. The cascaded coefficients (95% CI) from digital access to digital use, then to mHealth use, and then to time and satisfaction with health care were 0.28 (95% CI 0.22-0.35), 0.51 (95% CI 0.38-0.64), and 0.14 (95% CI 0.05-0.22), respectively. mHealth use, however, had a weakly significant effect on time and satisfaction with health care. mHealth use was not significantly associated with

sociodemographic variables, except for employment status, with which it had a weak correlation.

The potential influence of sociodemographic characteristics reflected the level of digital divide. The number of variables

and the magnitude of the coefficients were higher for digital use than for digital access and mHealth use. Thus, use divide in our setting was the most important gap.

Table 3. Correlation analysis (Pearson r and 2-tailed P value) among latent variables by confirmatory factor analysis.

Latent variable	Digital access	Digital use	mHealth ^a use	Time and satisfaction with health care
Digital access				
r	1.000	0.718 ^b	0.417 ^b	0.159 ^b
P value	— ^c	<.001	<.001	.001
Digital use				
r	0.718 ^b	1.000	0.607 ^b	0.226 ^b
P value	<.001	—	<.001	<.001
mHealth use				
r	0.417 ^b	0.607 ^b	1.000	0.231 ^b
P value	<.001	<.001	—	<.001
Time and satisfaction with health care				
r	0.159 ^b	0.226 ^b	0.231 ^b	1.000
P value	.001	<.001	<.001	—
Cronbach α	.745	.912	.857	.835
Average variance extracted	0.582	0.639	0.559	0.587

^amHealth: mobile health.

^bThe correlation is significant at a significance value of .05 (2-tailed).

^cNot applicable.

Figure 3. Structural equation modeling for digital divide in mobile health (mHealth). Solid lines represent significant relationships, and dotted lines represent nonsignificant ones; numbers on the lines from sociodemographic variables to latent variables are standardized coefficients, and numbers on the lines from latent variables to items are loadings. *** $P < .001$, ** $P < .01$, and * $P < .05$.

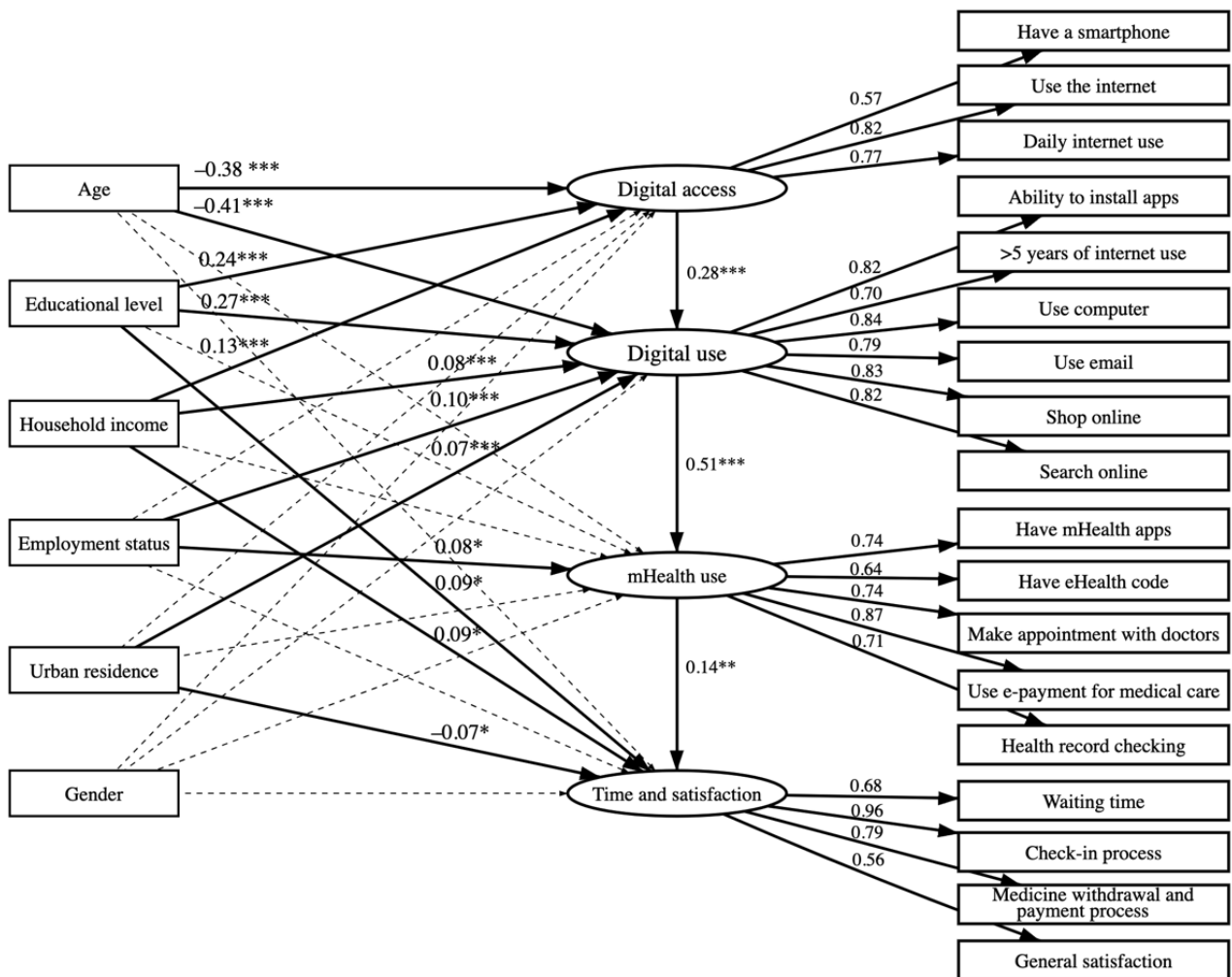


Table 4. Regression weights of parameters by the multiple-indicator, multiple-cause model with structural equation modeling.

Link	β coefficient (95% CI)	P value
Age → digital access	-0.38 (-0.45 to -0.31)	<.001
Age → digital use	-0.41 (-0.46 to -0.36)	<.001
Age → mHealth ^a use	-0.04 (-0.15 to 0.06)	.40
Age → time and satisfaction with health care	-0.08 (-0.16 to 0.01)	.07
Educational level → digital access	0.24 (0.15 to 0.32)	<.001
Educational level → digital use	0.27 (0.22 to 0.33)	<.001
Educational level → mHealth use	0.04 (-0.05 to 0.12)	.41
Educational level → time and satisfaction with health care	0.09 (0.00 to 0.17)	.04
Household income → digital access	0.13 (0.06 to 0.19)	<.001
Household income → digital use	0.08 (0.04 to 0.12)	<.001
Household income → mHealth use	0.02 (-0.05 to 0.08)	.62
Household income → time and satisfaction with health care	0.09 (0.02 to 0.17)	.01
Employment status → digital access	0.03 (-0.02 to 0.09)	.20
Employment status → digital use	0.10 (0.07 to 0.14)	<.001
Employment status → mHealth use	0.08 (0.02 to 0.15)	.01
Employment status → time and satisfaction with health care	0.003 (-0.07 to 0.07)	.94
Urban residence → digital access	-0.02 (-0.09 to 0.05)	.53
Urban residence → digital use	0.07 (0.03 to 0.11)	<.001
Urban residence → mHealth use	-0.04 (-0.09 to 0.02)	.20
Urban residence → time and satisfaction with health care	-0.07 (-0.14 to -0.01)	.03
Gender → digital access	0.06 (0.00 to 0.11)	.06
Gender → digital use	-0.01 (-0.05 to 0.02)	.41
Gender → mHealth use	-0.05 (-0.11 to 0.00)	.07
Gender → time and satisfaction with health care	0.006 (-0.06 to 0.07)	.85
Digital access → digital use	0.28 (0.22 to 0.35)	<.001
Digital use → mHealth use	0.51 (0.38 to 0.64)	<.001
mHealth use → time and satisfaction with health care	0.14 (0.05 to 0.22)	.002

^amHealth: mobile health.

Discussion

Principal Findings

We confirmed the framework of digital divide in mHealth app use. Our subjects were mostly educated and middle class, with good experience in internet use and other digital media. Between one-half to two-thirds were using basic mHealth features. Sociodemographic factors had stronger potential influences on digital use than on digital access; they also had the least direct effect on mHealth use and time and satisfaction with health care. However, mHealth use was potentially influenced by digital use. Time and satisfaction with health care, on the other hand, was only weakly associated with mHealth use.

The data from this study identified a cascade of potential successive influences, where digital access influenced digital use, which then influenced mHealth use. mHealth use was determined by digital access and use. Similar to our study, a

study by Tirado-Morueta et al [20] found that there was an indirect potential influence pathway from physical access to operative use and expressive informative use of the internet; ignoring this intermediary and simple running regression that predicted mHealth use from sociodemographics would lead to a misinterpretation of the results.

Among the sociodemographic variables examined, age and educational level were the stronger potential influencing variables. Both had direct independent influence on digital access and use, but had no direct effect on mHealth use; in addition, educational level had little effect on time and satisfaction with health care. Thus, their effects occurred in the early part of the digital chain. Based on this potential pathway, assistance for older adults and less educated public clients would need to start with improvements in access and use of the internet in general, as well as mobile facilities, such as email, social

media, and online business. Experience with these will make mHealth use easier for them.

Sociodemographic variables had a stronger potential influence on digital use than on digital access. This may reflect that the use divide was due to lifestyle differences more than it was due to inequity problems [56]. The cost of digital access in China was relatively small (around US \$15/60 Mbps or more per month [57]) and, hence, did not contribute much to digital access inequalities. On the other hand, certain sociodemographic groups, such as youth and upper-class people, choosing or needing to use digital technology is due to their lifestyle [58]. The small but significant correlation ($\beta=0.28$) between digital access and digital use may, in fact, reflect a noncausal relationship.

With the use of SEM, our findings on the effects of sociodemographic variables were different from those of other studies using one-step regression [59,60]. These other studies showed that sociodemographic variables potentially influenced mHealth use, but they missed the fact that the effect passed through digital use. Their findings would indicate the emphasis to improve mHealth use among the underprivileged. Our findings, on the other hand, imply that improved general digital use would be a more natural way to empower these groups of clients. This will make it easier and probably more effective to introduce mHealth to them. To reduce the existing digital divide among hospital clients, the hospital administration should provide special services or appropriate education to assist clients in making better use of mHealth apps.

According to another study in China, mHealth was effective in reducing patient waiting times and increasing patient satisfaction

in tertiary hospitals [26]. Another study found that waiting times for consultations and prescription filling reduced, resulting in increased outpatient satisfaction with pharmacy services [27]. Moreover, our study validated the marginally significant effect of mHealth use on shorter waiting times and improved satisfaction. This indicates that mHealth app use cannot adequately explain shorter waiting times and increased satisfaction. This indicates a need for further study. Additionally, since mHealth app use in hospitals is in its infancy, mHealth apps must be improved in terms of design and marketing based on existing digital use to increase their use and provide benefits to clients [33].

Limitations

This was a cross-sectional study. One may argue that the causation proposed in the SEM was limited by temporal sequence and may not be valid. We argue that sociodemographic variables are long-term values and do not vary much over time, whereas digital divide in these domains only comes after. Mobile apps were developed nearly a decade after the wide use of the internet began, and hospital-based mHealth is the most recent development. Therefore, our proposed pathway may not be flawed in terms of temporal sequence. The current stage of mHealth development in our setting is changing continuously. Thus, further studies may produce different results.

Conclusions

In order to close the mHealth use divide, it is important to improve digital access and digital use among older adults and less educated groups.

Acknowledgments

Funding was received from the Higher Education Research Promotion and Thailand's Education Hub for the Southern Region of ASEAN (Association of Southeast Asian Nations) Countries Project Office of the Higher Education Commission and Inner Mongolia Medical University, China. We would like to express our thanks to the Information technology (IT) staff and all participants from the study hospitals. This study is part of the first author's thesis to fulfil the requirement for the degree of Doctor of Philosophy in epidemiology.

Conflicts of Interest

None declared.

References

1. Istepanian R, Lactal J. Emerging mobile communication technologies for health: Some imperative notes on m-health. In: Proceedings of the 25th Annual International Conference of the IEEE Engineering in Medicine and Biology Society. 2003 Presented at: The 25th Annual International Conference of the IEEE Engineering in Medicine and Biology Society; September 17-21, 2003; Cancun, Mexico p. 1414-1416. [doi: [10.1109/iembs.2003.1279581](https://doi.org/10.1109/iembs.2003.1279581)]
2. mHealth: New horizons for Health Through Mobile Technologies. Geneva, Switzerland: World Health Organization; 2011. URL: https://apps.who.int/iris/bitstream/handle/10665/44607/9789241564250_eng.pdf [accessed 2019-10-01]
3. Becker S, Miron-Shatz T, Schumacher N, Krocza J, Diamantidis C, Albrecht U. mHealth 2.0: Experiences, possibilities, and perspectives. JMIR Mhealth Uhealth 2014 May 16;2(2):e24 [FREE Full text] [doi: [10.2196/mhealth.3328](https://doi.org/10.2196/mhealth.3328)] [Medline: [25099752](https://pubmed.ncbi.nlm.nih.gov/25099752/)]
4. 2019-2020 Investigation Report on the State of Informatization of Chinese Hospitals. Beijing, China: China Hospital Information Management Association; 2021. URL: <https://www.chima.org.cn/Html/News/Articles/8688.html> [accessed 2021-06-01]

5. Liu H, Lee W, Sun Y, Fen J, Chen T, Chou L, et al. Hospital-owned apps in Taiwan: Nationwide survey. *JMIR Mhealth Uhealth* 2018 Jan 16;6(1):e22 [FREE Full text] [doi: [10.2196/mhealth.8636](https://doi.org/10.2196/mhealth.8636)] [Medline: [29339347](https://pubmed.ncbi.nlm.nih.gov/29339347/)]
6. Ming LC, Untong N, Aliudin NA, Osili N, Kifli N, Tan CS, et al. Mobile health apps on COVID-19 launched in the early days of the pandemic: Content analysis and review. *JMIR Mhealth Uhealth* 2020 Sep 16;8(9):e19796 [FREE Full text] [doi: [10.2196/19796](https://doi.org/10.2196/19796)] [Medline: [32609622](https://pubmed.ncbi.nlm.nih.gov/32609622/)]
7. Grill E, Eitze S, De Bock F, Draganó N, Huebl L, Schmich P, et al. Sociodemographic characteristics determine download and use of a corona contact tracing app in Germany-Results of the COSMO surveys. *PLoS One* 2021;16(9):e0256660 [FREE Full text] [doi: [10.1371/journal.pone.0256660](https://doi.org/10.1371/journal.pone.0256660)] [Medline: [34473733](https://pubmed.ncbi.nlm.nih.gov/34473733/)]
8. Salathé M, Althaus CL, Neher R, Stringhini S, Hodcroft E, Fellay J, et al. COVID-19 epidemic in Switzerland: On the importance of testing, contact tracing and isolation. *Swiss Med Wkly* 2020 Mar 09;150:w20225 [FREE Full text] [doi: [10.4414/smw.2020.20225](https://doi.org/10.4414/smw.2020.20225)] [Medline: [32191813](https://pubmed.ncbi.nlm.nih.gov/32191813/)]
9. van Deursen AJ. Digital inequality during a pandemic: Quantitative study of differences in COVID-19-related internet uses and outcomes among the general population. *J Med Internet Res* 2020 Aug 20;22(8):e20073 [FREE Full text] [doi: [10.2196/20073](https://doi.org/10.2196/20073)] [Medline: [32750005](https://pubmed.ncbi.nlm.nih.gov/32750005/)]
10. Giansanti D, Veltro G. The digital divide in the era of COVID-19: An investigation into an important obstacle to the access to the by the citizen. *Healthcare (Basel)* 2021 Mar 26;9(4):371 [FREE Full text] [doi: [10.3390/healthcare9040371](https://doi.org/10.3390/healthcare9040371)] [Medline: [33810572](https://pubmed.ncbi.nlm.nih.gov/33810572/)]
11. Hannemann N, Götz NA, Schmidt L, Hübner U, Babitsch B. Patient connectivity with healthcare professionals and health insurer using digital health technologies during the COVID-19 pandemic: A German cross-sectional study. *BMC Med Inform Decis Mak* 2021 Aug 25;21(1):250 [FREE Full text] [doi: [10.1186/s12911-021-01605-8](https://doi.org/10.1186/s12911-021-01605-8)] [Medline: [34433452](https://pubmed.ncbi.nlm.nih.gov/34433452/)]
12. Yan A, Zou Y, Mirchandani DA. How hospitals in mainland China responded to the outbreak of COVID-19 using information technology-enabled services: An analysis of hospital news webpages. *J Am Med Inform Assoc* 2020 Jul 01;27(7):991-999 [FREE Full text] [doi: [10.1093/jamia/ocaa064](https://doi.org/10.1093/jamia/ocaa064)] [Medline: [32311036](https://pubmed.ncbi.nlm.nih.gov/32311036/)]
13. Gong K, Xu Z, Cai Z, Chen Y, Wang Z. Internet hospitals help prevent and control the epidemic of COVID-19 in China: Multicenter user profiling study. *J Med Internet Res* 2020 Apr 14;22(4):e18908 [FREE Full text] [doi: [10.2196/18908](https://doi.org/10.2196/18908)] [Medline: [32250962](https://pubmed.ncbi.nlm.nih.gov/32250962/)]
14. Wang Y, Liu Y, Shi Y, Yu Y, Yang J. User perceptions of virtual hospital apps in China: Systematic search. *JMIR Mhealth Uhealth* 2020 Aug 12;8(8):e19487 [FREE Full text] [doi: [10.2196/19487](https://doi.org/10.2196/19487)] [Medline: [32687480](https://pubmed.ncbi.nlm.nih.gov/32687480/)]
15. Falling through the net: Defining the digital divide. National Telecommunications and Information Administration, US Department of Commerce. 1999. URL: <https://www.ntia.doc.gov/legacy/ntiahome/ftn99/contents.html> [accessed 2019-12-04]
16. Compaine BM, editor. *The Digital Divide: Facing a Crisis or Creating a Myth?*. Cambridge, MA: The MIT Press; 2001:1-357.
17. van Dijk JAGM. Digital divide research, achievements and shortcomings. *Poetics* 2006 Aug;34(4-5):221-235. [doi: [10.1016/j.poetic.2006.05.004](https://doi.org/10.1016/j.poetic.2006.05.004)]
18. Srinuan C, Bohlin E. Understanding the digital divide: A literature survey and ways forward. In: Proceedings of the 22nd European Regional Conference of the International Telecommunications Society. 2011 Presented at: The 22nd European Regional Conference of the International Telecommunications Society; September 18-21, 2011; Budapest, Hungary p. 1-38 URL: <https://www.econstor.eu/bitstream/10419/52191/1/672623358.pdf>
19. Wei K, Teo H, Chan HC, Tan BCY. Conceptualizing and testing a social cognitive model of the digital divide. *Inf Syst Res* 2011 Mar;22(1):170-187. [doi: [10.1287/isre.1090.0273](https://doi.org/10.1287/isre.1090.0273)]
20. Tirado-Morueta R, Aguaded-Gómez J, Hernando-Gómez Á. The socio-demographic divide in internet usage moderated by digital literacy support. *Technol Soc* 2018 Nov;55:47-55. [doi: [10.1016/j.techsoc.2018.06.001](https://doi.org/10.1016/j.techsoc.2018.06.001)]
21. The 47th Statistical Report on China's Internet Development. Beijing, China: China Internet Network Information Center; 2021 Feb. URL: <https://www.cnnic.com.cn/IDR/ReportDownloads/202104/P020210420557302172744.pdf> [accessed 2021-10-06]
22. Zhao D, Zhao H, Cleary PD. Understanding the determinants of public trust in the health care system in China: An analysis of a cross-sectional survey. *J Health Serv Res Policy* 2019 Jan 03;24(1):37-43. [doi: [10.1177/1355819618799113](https://doi.org/10.1177/1355819618799113)] [Medline: [30176742](https://pubmed.ncbi.nlm.nih.gov/30176742/)]
23. Prakash B. Patient satisfaction. *J Cutan Aesthet Surg* 2010 Sep;3(3):151-155 [FREE Full text] [doi: [10.4103/0974-2077.74491](https://doi.org/10.4103/0974-2077.74491)] [Medline: [21430827](https://pubmed.ncbi.nlm.nih.gov/21430827/)]
24. Hussain A, Asif M, Jameel A, Hwang J. Measuring OPD patient satisfaction with different service delivery aspects at public hospitals in Pakistan. *Int J Environ Res Public Health* 2019 Jul 02;16(13):2340 [FREE Full text] [doi: [10.3390/ijerph16132340](https://doi.org/10.3390/ijerph16132340)] [Medline: [31269722](https://pubmed.ncbi.nlm.nih.gov/31269722/)]
25. Gimpel H, Manner-Romberg T, Schmied F, Winkler TJ. Understanding the evaluation of mHealth app features based on a cross-country Kano analysis. *Electron Mark* 2021 Mar 24;31(4):765-794. [doi: [10.1007/s12525-020-00455-y](https://doi.org/10.1007/s12525-020-00455-y)]
26. Xie W, Yang X, Cao X, Liu P. Effects of a comprehensive reservation service for non-emergency registration on appointment registration rate, patient waiting time, patient satisfaction and outpatient volume in a tertiary hospital in China. *BMC Health Serv Res* 2019 Nov 01;19(1):782 [FREE Full text] [doi: [10.1186/s12913-019-4652-6](https://doi.org/10.1186/s12913-019-4652-6)] [Medline: [31675973](https://pubmed.ncbi.nlm.nih.gov/31675973/)]

27. Sun J, Lin Q, Zhao P, Zhang Q, Xu K, Chen H, et al. Reducing waiting time and raising outpatient satisfaction in a Chinese public tertiary general hospital-An interrupted time series study. *BMC Public Health* 2017 Aug 22;17(1):668 [FREE Full text] [doi: [10.1186/s12889-017-4667-z](https://doi.org/10.1186/s12889-017-4667-z)] [Medline: [28830400](https://pubmed.ncbi.nlm.nih.gov/28830400/)]
28. Sun J, Hu G, Ma J, Chen Y, Wu L, Liu Q, et al. Consumer satisfaction with tertiary healthcare in China: Findings from the 2015 China National Patient Survey. *Int J Qual Health Care* 2017 Apr 01;29(2):213-221. [doi: [10.1093/intqhc/mzw160](https://doi.org/10.1093/intqhc/mzw160)] [Medline: [28096280](https://pubmed.ncbi.nlm.nih.gov/28096280/)]
29. Lu C, Hu Y, Xie J, Fu Q, Leigh I, Governor S, et al. The use of mobile health applications to improve patient experience: Cross-sectional study in Chinese public hospitals. *JMIR Mhealth Uhealth* 2018 May 23;6(5):e126 [FREE Full text] [doi: [10.2196/mhealth.9145](https://doi.org/10.2196/mhealth.9145)] [Medline: [29792290](https://pubmed.ncbi.nlm.nih.gov/29792290/)]
30. Xie W, Cao X, Dong H, Liu Y. The use of smartphone-based triage to reduce the rate of outpatient error registration: Cross-sectional study. *JMIR Mhealth Uhealth* 2019 Nov 11;7(11):e15313 [FREE Full text] [doi: [10.2196/15313](https://doi.org/10.2196/15313)] [Medline: [31710300](https://pubmed.ncbi.nlm.nih.gov/31710300/)]
31. Ye Q, Deng Z, Chen Y, Liao J, Li G, Lu Y. How resource scarcity and accessibility affect patients' usage of mobile health in China: Resource competition perspective. *JMIR Mhealth Uhealth* 2019 Aug 09;7(8):e13491 [FREE Full text] [doi: [10.2196/13491](https://doi.org/10.2196/13491)] [Medline: [31400104](https://pubmed.ncbi.nlm.nih.gov/31400104/)]
32. Shen L, Wang S, Chen W, Fu Q, Evans R, Lan F, et al. Understanding the function constitution and influence factors on communication for the WeChat official account of top tertiary hospitals in China: Cross-sectional study. *J Med Internet Res* 2019 Dec 09;21(12):e13025 [FREE Full text] [doi: [10.2196/13025](https://doi.org/10.2196/13025)] [Medline: [31815674](https://pubmed.ncbi.nlm.nih.gov/31815674/)]
33. Yu W, Yu X, Hu H, Duan G, Liu Z, Wang Y. Use of hospital appointment registration systems in China: A survey study. *Glob J Health Sci* 2013 Jul 22;5(5):193-201 [FREE Full text] [doi: [10.5539/gjhs.v5n5p193](https://doi.org/10.5539/gjhs.v5n5p193)] [Medline: [23985121](https://pubmed.ncbi.nlm.nih.gov/23985121/)]
34. Sieck CJ, Sheon A, Ancker JS, Castek J, Callahan B, Siefer A. Digital inclusion as a social determinant of health. *NPJ Digit Med* 2021 Mar 17;4(1):52 [FREE Full text] [doi: [10.1038/s41746-021-00413-8](https://doi.org/10.1038/s41746-021-00413-8)] [Medline: [33731887](https://pubmed.ncbi.nlm.nih.gov/33731887/)]
35. Free C, Phillips G, Felix L, Galli L, Patel V, Edwards P. The effectiveness of M-health technologies for improving health and health services: A systematic review protocol. *BMC Res Notes* 2010 Oct 06;3:250 [FREE Full text] [doi: [10.1186/1756-0500-3-250](https://doi.org/10.1186/1756-0500-3-250)] [Medline: [20925916](https://pubmed.ncbi.nlm.nih.gov/20925916/)]
36. Be He@lthy, Be Mobile: Annual Report 2017. Geneva, Switzerland: World Health Organization and International Telecommunication Union; 2018. URL: <https://apps.who.int/iris/bitstream/handle/10665/326498/9789241516242-eng.pdf?sequence=1&isAllowed=y> [accessed 2022-05-17]
37. Hamilton EC, Saiyed F, Miller CC, Eguia A, Fonseca AC, Baum GP, et al. The digital divide in adoption and use of mobile health technology among caregivers of pediatric surgery patients. *J Pediatr Surg* 2018 Aug;53(8):1478-1493. [doi: [10.1016/j.jpedsurg.2017.08.023](https://doi.org/10.1016/j.jpedsurg.2017.08.023)] [Medline: [28927983](https://pubmed.ncbi.nlm.nih.gov/28927983/)]
38. Wang L, Xiao Y, Rao E, Jiang L, Xiao Y, Ouyang Z. An assessment of the impact of urbanization on soil erosion in Inner Mongolia. *Int J Environ Res Public Health* 2018 Mar 19;15(3):550 [FREE Full text] [doi: [10.3390/ijerph15030550](https://doi.org/10.3390/ijerph15030550)] [Medline: [29562707](https://pubmed.ncbi.nlm.nih.gov/29562707/)]
39. Li M, Fan Y, McNeil EB, Chongsuvivatwong V. Traditional Mongolian, Traditional Chinese, and Western Medicine hospitals: System review and patient survey on expectations and perceptions of quality of healthcare in Inner Mongolia, China. *Evid Based Complement Alternat Med* 2018;2018:2698461 [FREE Full text] [doi: [10.1155/2018/2698461](https://doi.org/10.1155/2018/2698461)] [Medline: [30108650](https://pubmed.ncbi.nlm.nih.gov/30108650/)]
40. Hu Y, Zhang Z, Xie J, Wang G. The Outpatient Experience Questionnaire of comprehensive public hospital in China: Development, validity and reliability. *Int J Qual Health Care* 2017 Feb 01;29(1):40-46. [doi: [10.1093/intqhc/mzw133](https://doi.org/10.1093/intqhc/mzw133)] [Medline: [27836998](https://pubmed.ncbi.nlm.nih.gov/27836998/)]
41. Chang C, Gardiner J, Houang R, Yu Y. Comparing multiple statistical software for multiple-indicator, multiple-cause modeling: An application of gender disparity in adult cognitive functioning using MIDUS II dataset. *BMC Med Res Methodol* 2020 Nov 12;20(1):275 [FREE Full text] [doi: [10.1186/s12874-020-01150-4](https://doi.org/10.1186/s12874-020-01150-4)] [Medline: [33183226](https://pubmed.ncbi.nlm.nih.gov/33183226/)]
42. Fox J, Dusa A. Package 'polycor': Polychoric and polyserial correlations. The Comprehensive R Archive Network. 2022 Jan 11. URL: <https://cran.r-project.org/web/packages/polycor/polycor.pdf> [accessed 2022-05-07]
43. Revelle W. Package 'psych': Procedures for psychological, psychometric, and personality research. Personality Project. 2021 Mar 27. URL: <https://personality-project.org/r/psych-manual.pdf> [accessed 2022-03-01]
44. Yves RT, Jorgensen TD, Rockwood N, Oberski D, Byrnes J, Vanbrabant L, et al. Package 'lavaan': Latent variable analysis. The Comprehensive R Archive Network. 2022 Mar 31. URL: <https://cran.r-project.org/web/packages/lavaan/lavaan.pdf> [accessed 2022-04-05]
45. Xie Z, Nacioglu A, Or C. Prevalence, demographic correlates, and perceived impacts of mobile health app use amongst Chinese adults: Cross-sectional survey study. *JMIR Mhealth Uhealth* 2018 Apr 26;6(4):e103 [FREE Full text] [doi: [10.2196/mhealth.9002](https://doi.org/10.2196/mhealth.9002)] [Medline: [29699971](https://pubmed.ncbi.nlm.nih.gov/29699971/)]
46. Households' income and consumption expenditure in the first half year of 2021. National Bureau of Statistics of China. 2021 Jul 16. URL: http://www.stats.gov.cn/enGliSH/PressRelease/202107/t20210716_1819567.html [accessed 2022-05-10]
47. Kaiser HF. A second generation Little Jiffy. *Psychometrika* 1970 Dec;35(4):401-415. [doi: [10.1007/BF02291817](https://doi.org/10.1007/BF02291817)]
48. Kaiser HF, Rice J. Little Jiffy, Mark IV. *Educ Psychol Meas* 2016 Jul 02;34(1):111-117. [doi: [10.1177/001316447403400115](https://doi.org/10.1177/001316447403400115)]

49. Bartlett MS. The effect of standardization on a χ^2 approximation in factor analysis. *Biometrika* 1951 Dec;38(3/4):337. [doi: [10.2307/2332580](https://doi.org/10.2307/2332580)]
50. Courtney MGR. Determining the number of factors to retain in EFA: Using the SPSS R-Menu v2.0 to make more judicious estimations. *Pract Assess Res Eval* 2013;18:1-14 [FREE Full text] [doi: [10.7275/9cf5-2m72](https://doi.org/10.7275/9cf5-2m72)]
51. Brown TA. *Confirmatory Factor Analysis for Applied Research*. 2nd edition. New York, NY: The Guilford Press; 2015.
52. Fabrigar LR, Wegener DT. *Exploratory Factor Analysis: Understanding Statistics*. New York, NY: Oxford University Press; 2011.
53. Hair Jr JF, Black WC, Babin BJ, Anderson RE. *Multivariate Data Analysis: A Global Perspective*. 7th edition. London, UK: Pearson Education; 2010.
54. Pettersson E, Turkheimer E. Item selection, evaluation, and simple structure in personality data. *J Res Pers* 2010 Aug 01;44(4):407-420 [FREE Full text] [doi: [10.1016/j.jrp.2010.03.002](https://doi.org/10.1016/j.jrp.2010.03.002)] [Medline: [20694168](https://pubmed.ncbi.nlm.nih.gov/20694168/)]
55. Gana K, Broc G. *Structural Equation Modeling with lavaan*. London, UK: ISTE Ltd; 2019.
56. Hofmann G, Thomas G. Digital lifestyle 2020. *IEEE Multimed* 2008 Apr;15(2):4-7. [doi: [10.1109/mmul.2008.24](https://doi.org/10.1109/mmul.2008.24)]
57. Price rankings by country of internet (60 Mbps or more, unlimited data, cable/ADSL) (utilities (monthly)). Numbeo. 2022. URL: https://www.numbeo.com/cost-of-living/country_price_rankings?itemId=33 [accessed 2021-11-01]
58. Winkelhake U. "Digital lifestyle"—Future employees and customers. In: *The Digital Transformation of the Automotive Industry*. Cham, Switzerland: Springer; 2021:29-46.
59. Carroll JK, Moorhead A, Bond R, LeBlanc WG, Petrella RJ, Fiscella K. Who uses mobile phone health apps and does use matter? A secondary data analytics approach. *J Med Internet Res* 2017 Apr 19;19(4):e125 [FREE Full text] [doi: [10.2196/jmir.5604](https://doi.org/10.2196/jmir.5604)] [Medline: [28428170](https://pubmed.ncbi.nlm.nih.gov/28428170/)]
60. Ernsting C, Dombrowski SU, Oedekoven M, O'Sullivan JL, Kanzler M, Kuhlmeier A, et al. Using smartphones and health apps to change and manage health behaviors: A population-based survey. *J Med Internet Res* 2017 Apr 05;19(4):e101 [FREE Full text] [doi: [10.2196/jmir.6838](https://doi.org/10.2196/jmir.6838)] [Medline: [28381394](https://pubmed.ncbi.nlm.nih.gov/28381394/)]

Abbreviations

ASEAN: Association of Southeast Asian Nations
CFA: confirmatory factor analysis
CFI: comparative fit index
EFA: exploratory factor analysis
e-payment: electronic payment
mHealth: mobile health
MIMIC: multiple-indicator, multiple-cause model
RMSEA: root mean square error of approximation
SEM: structural equation modeling
SRMR: standardized root mean square residual
TLI: Tucker-Lewis index

Edited by G Eysenbach, A Kushniruk; submitted 01.02.22; peer-reviewed by Y Wang; comments to author 15.03.22; revised version received 20.03.22; accepted 22.04.22; published 19.05.22.

Please cite as:

Cao L, Chongsuvivatwong V, McNeil EB

The Sociodemographic Digital Divide in Mobile Health App Use Among Clients at Outpatient Departments in Inner Mongolia, China: Cross-sectional Survey Study

JMIR Hum Factors 2022;9(2):e36962

URL: <https://humanfactors.jmir.org/2022/2/e36962>

doi: [10.2196/36962](https://doi.org/10.2196/36962)

PMID: [35587367](https://pubmed.ncbi.nlm.nih.gov/35587367/)

©Li Cao, Virasakdi Chongsuvivatwong, Edward B McNeil. Originally published in *JMIR Human Factors* (<https://humanfactors.jmir.org>), 19.05.2022. This is an open-access article distributed under the terms of the Creative Commons Attribution License (<https://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in *JMIR Human Factors*, is properly cited. The complete bibliographic information, a link to the original publication on <https://humanfactors.jmir.org>, as well as this copyright and license information must be included.

Original Paper

Understanding People With Chronic Pain Who Use a Cognitive Behavioral Therapy–Based Artificial Intelligence Mental Health App (Wysa): Mixed Methods Retrospective Observational Study

Saha Meheli^{1*}, MA, MPhil; Chaitali Sinha^{2*}, MA; Madhura Kadaba^{2*}, MS

¹Department of Clinical Psychology, National Institute of Mental Health and Neurosciences, Bangalore, India

²Wysa Inc, Boston, MA, United States

* all authors contributed equally

Corresponding Author:

Chaitali Sinha, MA

Wysa Inc

131 Dartmouth St

Boston, MA

United States

Phone: 1 916 753 7824

Email: chaitali@wysa.io

Abstract

Background: Digital health interventions can bridge barriers in access to treatment among individuals with chronic pain.

Objective: This study aimed to evaluate the perceived needs, engagement, and effectiveness of the mental health app Wysa with regard to mental health outcomes among real-world users who reported chronic pain and engaged with the app for support.

Methods: Real-world data from users (N=2194) who reported chronic pain and associated health conditions in their conversations with the mental health app were examined using a mixed methods retrospective observational study. An inductive thematic analysis was used to analyze the conversational data of users with chronic pain to assess perceived needs, along with comparative macro-analyses of conversational flows to capture engagement within the app. Additionally, the scores from a subset of users who completed a set of pre-post assessment questionnaires, namely Patient Health Questionnaire-9 (PHQ-9) (n=69) and Generalized Anxiety Disorder Assessment-7 (GAD-7) (n=57), were examined to evaluate the effectiveness of Wysa in providing support for mental health concerns among those managing chronic pain.

Results: The themes emerging from the conversations of users with chronic pain included *health concerns*, *socioeconomic concerns*, and *pain management concerns*. Findings from the quantitative analysis indicated that users with chronic pain showed significantly greater app engagement ($P<.001$) than users without chronic pain, with a large effect size (Vargha and Delaney $A=0.76-0.80$). Furthermore, users with pre-post assessments during the study period were found to have significant improvements in group means for both PHQ-9 and GAD-7 symptom scores, with a medium effect size (Cohen $d=0.60-0.61$).

Conclusions: The findings indicate that users look for tools that can help them address their concerns related to mental health, pain management, and sleep issues. The study findings also indicate the breadth of the needs of users with chronic pain and the lack of support structures, and suggest that Wysa can provide effective support to bridge the gap.

(*JMIR Hum Factors* 2022;9(2):e35671) doi:[10.2196/35671](https://doi.org/10.2196/35671)

KEYWORDS

chronic pain; digital mental health; mobile health; mHealth; pain management; artificial intelligence; cognitive behavioral therapy; conversational agent; software agent; pain conditions; depression; anxiety

Introduction

Background

Chronic pain is a debilitating condition that affects over 2 billion people worldwide, with the global weighted mean prevalence

estimated to be 30.3% [1,2]. The sequelae of chronic pain are spread across the different aspects of an individual's life, such as employment, sleep, physical and mental health, and social relationships [3].

The relationship between chronic pain and mental health comorbidities is well documented, with the prevalence of concurrent major depression or anxiety diagnoses ranging between 35% and 50% of the global population, and the presence of pain often doubling the risk of having a mental health diagnosis [4,5]. Mental health concerns can compound the difficulties in gaining access to affordable and accessible treatment for chronic pain. The absence of referral pathways, high costs, and stigma increase barriers to access for treatment and management [6-8]. Increasingly, digital health care apps have been addressing these challenges with cost-effective and accessible solutions, with some evidence to indicate feasibility and efficacy [9,10]; however, adoption, acceptability, and quality still remain challenges for this space [11]. Currently, there is limited user-led research that captures the perspectives and perceived needs (based on the person's own judgement) of individuals with chronic pain.

This study examined the experiences of users with self-reported chronic pain who were using Wysa, which is an anonymous artificial intelligence (AI) conversational agent app for digital mental health. A previous study showed that the use of Wysa was associated with a significant reduction in depressive symptoms in a high-engagement user group compared with a low-engagement user group [12]. The Wysa app uses a free-text conversational interface to listen and respond to the user and responds to the user's distress by recommending evidence-based elements from cognitive behavioral therapy (CBT), behavioral reinforcement, and mindfulness, among others. The self-help practices and conversation-based tools and techniques provide support for challenges, including anxiety, sleep, low energy, motivation, loss, and pain.

Our Study

The purpose of this study was to understand the perceived needs of individuals with chronic pain and their mental health concerns. The study aimed to understand the ways in which individuals find digital health support meaningful, and the potential areas in which it could be developed in order to offer more effective and accessible support for individuals with chronic pain.

This study had the following 3 objectives: (1) to evaluate the perceived needs of users with chronic pain conditions; (2) to evaluate the app engagement and disengagement patterns of users with chronic pain (in terms of the most used tools, most frequented conversational flows, and intensity of engagement and disengagement); and (3) to evaluate the efficacy of the Wysa app and its interventions for improving mental health among a subsample of users who have completed 2 questionnaire assessments (ie, preintervention and postintervention) using the validated Patient Health Questionnaire-9 (PHQ-9) [13] and Generalized Anxiety Disorder Assessment-7 (GAD-7) [14].

Methods

Ethical Considerations

The Wysa app is publicly available as an app on the Android and iOS app stores. It has been designed to prioritize safety, privacy, and security by design. There is no user registration

required, and no personally identifiable information is asked at any time during app use. As the study involved analyzing real-world data from an anonymous nonclinical population, it was exempt from registration in a public trial registry (according to OHRP guidelines [15]). The users voluntarily downloaded the app after having consented to the app's Terms of Service and Privacy Policy. For ethical and privacy reasons, the authors did not have access to all the user messages. Only minimal and limited conversational data extracted based on keywords were used for this research, and no longitudinal data were used. The study data set was deidentified using one-way cryptographic functions. User data were adequately secured according to the organization's privacy, security, and safety policies. The study participants were informed about how they can exercise their rights to restrict processing of their data for research purposes.

Study Design

During the study period from October 2020 to October 2021, a total of 2194 users were identified to have reported chronic pain based on selected keywords in ongoing conversations with the conversational agent. We implemented a retrospective observational study with a mixed methods approach, given the objectives of the study and the nature of the data being analyzed. In a retrospective observational study, the sample is defined later and the data are already available [16]. This methodology has been used in multiple health care studies, such as those involving depression and anxiety [17-19], hypertension [20], postpartum mental health concerns [21], COVID-19-related concerns [22], and diabetes [23].

Outcome Measures and Data Types

The following measures and data types were used: (1) textual snippets from users, (2) tool usage data, (3) usage data indicating the interventions used with the bot, and (4) self-reported PHQ-9 and GAD-7 data.

Procedure

Data Extraction

To extract the data, the first step was to optimize the keywords. For this, the classification of chronic pain by the International Association for the Study of Pain was used as a guide [24] to derive keywords. Additionally, other relevant keywords were derived from literature on pain and the clinical experiences of the researchers in this study. As such, the keywords also included general terms used to describe pain or to describe pain-related experiences, such as "pain," "painkiller," "manage pain," "nerve block," etc.

User messages that had at least one or more of the keywords were extracted, which resulted in 83,000 messages. The criteria for data extraction and filtration are listed in [Textbox 1](#). These were further filtered by the research team, who read through each of the messages to exclude any pain that did not relate to chronic pain conditions (such as emotional pain or menstrual pain) or pain that the users did not experience themselves. Ambiguous messages that did not clarify the nature of the pain were also excluded. For example, the word "pain" often created false positives or ambiguity, as users could use this word in a context that would not necessarily be related to physical pain

or chronic pain (eg, “it is so painful”). These types of messages were excluded from the final set. The stringent exclusion resulted in a final set of 3300 relevant messages from 2194

unique users. Among these, the distribution of the number of mentions of different diagnoses is listed in [Table 1](#).

Textbox 1. Criteria and keywords for data extraction.

Selection criteria (inclusion and exclusion) for data extraction

1. The keywords used were about “chronic pain” or related concerns to capture all mentions of associated conditions, such as “disability,” “loss of limb,” “manage pain,” “injury,” “musculoskeletal,” “neuralgia,” “back pain,” “phantom pain,” “multiple sclerosis,” “osteoarthritis,” “cancer pain,” “sciatica,” “fibromyalgia,” “spinal cord,” “arthritis,” “spondylitis,” “cervical,” “rheumatoid,” “endometriosis,” “surgery,” “ankylosing,” “inflammation,” “inflammatory,” “crohn,” “sclerosis,” “gout,” “nerve block,” “analgesic,” “painkillers,” “pelvic pain,” “cyst,” “migraine,” and “pcos.”
2. The report of pain should be first-person, that is, the person is talking about their own pain. Any mention of others’ health conditions was excluded.
3. There was clarity about the mentioned pain being chronic and physical in nature. Ambiguous messages that did not meet this criterion were excluded.
4. Any message that qualified only as emotional or psychosomatic pain was excluded.

Table 1. The number of mentions of different conditions of chronic pain.

Condition	Number of mentions ^a
Migraine	551
Back pain	501
Chronic pain	265
Surgery	212
Fibromyalgia	119
Injury	99
Disability	92
Neck pain	56
Arthritis	30
Crohn	22
Sciatica	19
Inflammation	18
Cervical	9
Rheumatoid	9
Sclerosis	9
Endometriosis	8
Spondylitis	5
Cyst/polycystic ovary syndrome	5
Multiple sclerosis	4
Nerve block	2
Phantom pain	1

^aThis is the number of actual mentions of the condition and not the number of users having the condition. Some users reported multiple conditions, and these categories are not mutually exclusive.

Analyses

Objective 1: Perceived Needs of Users With Chronic Pain

Inductive thematic analysis [25] was used to gather information on perceived needs and experiences related to chronic pain. The author SM got familiarized with the data by reading the

messages multiple times. This was followed by the generation of preliminary codes, which were then grouped into potential subthemes and themes. The data were then verified for relevance to the respective themes at each level, and the initial set of themes was verified in relation to the coded extracts. The themes were selected and finalized based on their relevance to the objectives of the study and their salience in capturing perceived

needs. Though the analysis was led by SM, all the authors met at regular intervals to ensure that the themes were internally consistent and unique, and answered the research question. The themes and subthemes were then reviewed and finalized.

Objective 2: App Engagement and Disengagement of Users With Chronic Pain

The engagement of users was evaluated at the following 3 different levels: (1) the frequency of usage within Wysa, (2) the intensity of user engagement, and (3) the points of disengagement. Engagement refers to the number of initiated and completed interactions within the app, while the points of disengagement are instances where users stop communicating or engaging further with the conversational agent.

The first level of analysis was the frequency count of the tools (interventions) most utilized by the users with chronic pain (N=2194). To capture the most recurring conversation flows of the users with chronic pain, a network analysis [26] was conducted with cleaned sequential pairs of conversational path units and their frequency of occurrence as weights. The path units were then grouped into higher units based on app elements, and the networks were visualized in Gephi 0.9.2 [27] as a directed weighted network. The directionality of the conversational pathways was then used to identify the tools most used within Wysa, to further validate the findings of the first-level analysis. The visualization of the network analysis was done in Gephi with the Yifan Hu layout, and the appearance of nodes and edges was based on “weighted in-degree,” that is, incoming connections as calculated within the software.

At the second level, to examine the intensity of engagement, the extracted data were analyzed for instances of engagement and disengagement with the bot. Using Python 3.6 code, the first-level engagement of a user was calculated by using the number of conversational pathways per user. The intensity of engagement and disengagement of users with chronic pain was then tested in comparison to users without chronic pain (a randomized sample from the larger user base that had not

reported any chronic pain). The data for both samples did not fit the criteria for normality, indicating the need for nonparametric assessments. The data were further tested using R 4.1, and the difference in engagement was assessed using the Mann-Whitney *U* test between users with chronic pain and those without chronic pain.

To evaluate the disengagement patterns, the ends of conversational flows, which represented the locations where users would stop engaging with the app, were identified for both groups. The top 20 conversational ends were located from path data and compared between the groups (users with chronic pain, n=2194; users without chronic pain, n=1880) for salient differences.

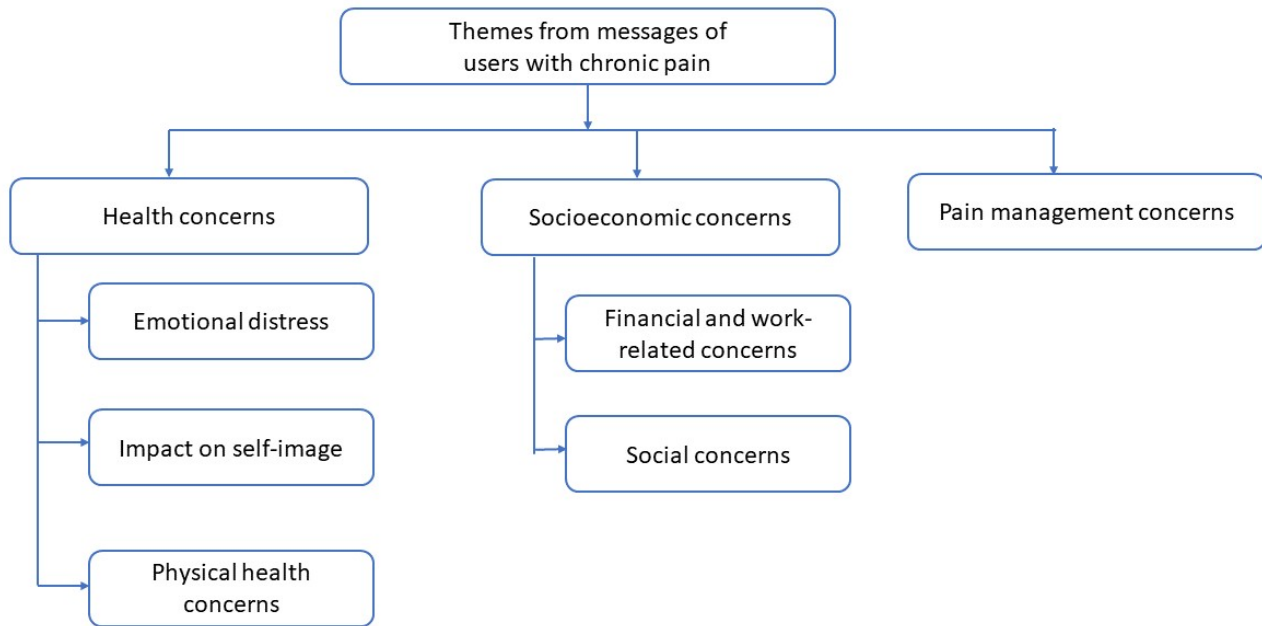
Objective 3: Efficacy in Improving the Mental Health of Users With Chronic Pain

A subset of the sample was used to examine the efficacy of the app for improving the mental health symptoms of users with chronic pain. This subset was restricted to users who completed at least two self-reported PHQ-9 (n=69) and GAD-7 (n=57) assessments during the study period. The first assessment during the study period was regarded as the baseline, and the last assessment was regarded as the postintervention assessment. This subsample was further restricted to users with a score greater than 5 on either the PHQ-9 or GAD-7 (mild distress or greater) at baseline. A paired *t* test was performed between the first and last assessments of the study period for this subset of users with chronic pain.

Results

Objective 1: Perceived Needs of Users With Chronic Pain

The thematic analysis yielded themes representative of the perceived needs and concerns of users with chronic pain. The themes (given in Figure 1) are *health concerns*, *socioeconomic concerns*, and *pain management concerns*.

Figure 1. Themes of messages from users with chronic pain (N=2194).

Health Concerns

This theme indicates the perceived needs, concerns, and experiences of users with regard to their mental and physical health. It includes the subthemes of *emotional distress*, *impact on self-image*, and *impact on physical health*.

Emotional Distress

The users with chronic pain would often write to Wysa about strong emotional difficulties, such as anxiety, feeling low or depressed, panic attacks, frustration, and irritability, among others. Many were often concerned with how chronic pain tended to increase their stress and affect their mental health. Some found themselves “flooded with suicidal thoughts and negative emotions.” This concern for the effect of chronic pain on one’s mental health was clearly demonstrated by a user who reported “it’s taking a big toll on my mental health.” The relationship between emotional distress and pain would often be cyclical or bidirectional, with some reporting how their stress levels often aggravated the pain, for example, “I have lupus, fibromyalgia, and asthma, and they are triggered by my stress levels.”

Impact on Self-Image

Chronic pain users often reported low self-confidence, with one user stating, “I am still young and having such a condition makes me feel inferior to other teens my age. I feel weak and useless” and others stating that they had “lower self-worth,” with the limitations impacting them deeply. A user said, “I have fibromyalgia, and it makes me feel like a failure at times.” Multiple users also reported having body image concerns, with some feeling “ugly” or feeling ashamed of how they appeared.

Impact on Physical Health

Apart from the significant pain experienced, this group of users also reported a variety of issues related to their physical health. Many indicated disturbed sleep (“I can never sleep well”) and

constant physical discomfort (“which means I can’t sit or stand or sleep properly”). They also worried about the possibility of their pain worsening and further physical distress. Some were “scared” of their upcoming surgeries. An often-reported issue was the limitations felt by users around physical activities and the resulting distress.

Socioeconomic Concerns

The experiences that users with chronic pain shared with Wysa included their experiences of their social relationships and their experiences related to work and finances. The subthemes under this theme are *financial and work-related concerns* and *social concerns*.

Financial and Work-Related Concerns

Users with chronic pain often reported several financial and work-related concerns. While some reported that they were unable to perform to their full potential due to chronic pain, others stated that chronic pain was “increasing work stress” and the possibility of “burnout.” The pain would be so debilitating at times that they would often find themselves unemployed, as highlighted by another user who stated, “made it hard for me to find work.” This work-related concern was also demonstrated in a response from a user.

I had to leave work early because of back pain. I feel inadequate for not being able to work with this body. It's frustrating and I feel like a burden.

This concern was further compounded by worries about the future, for example, a user wrote, “make(s) me worry if I will be able to survive in a capitalist system and if I will be able to sustain myself,” and financial worries about therapy and medical bills. Another user made the below statement.

Nothing I have tried for fibromyalgia has helped. I can't afford therapy. I can't afford food. I'm scared about everything all the time.

Social Concerns

Users with chronic pain often wrote to Wysa about how the pain limited their social interactions. This was clearly illustrated by a user who wrote, “couldn’t go out with my boyfriend to meet some friends” to portray how chronic pain could restrict someone’s mobility, preventing them from going out and socializing. Some users also reported how the pain and related health issues made them withdraw socially and reported feeling “lonely.” They also felt isolated in their pain and reported the feelings of not being understood, not receiving help in seeking professional help, and not being supported.

Well, I am in a relationship but seriously considering a split; he doesn't help, he's more of a hindrance. He argues with me when he's asked to help. I have chronic pain. He picks at the triggers as if it's a game.

Some also talked about the gratitude they felt for being supported by their family members (“Even though I had a terrible migraine today, my husband just made me feel like I was his whole world. I felt so loved”), which further highlighted the need for social support in individuals with chronic pain. This wish for validation and support was expressed by a user who wrote to Wysa as follows:

I am diagnosed with a bunch of things like compressed disk and pain in my back and I also have arthritis in both of my knees ... and a couple more things like neck problems. So, I just want everyone to say I understand how I feel, and everything will be okay and to just take it slow and take one step at a time and not rush or feel overwhelmed and tired.

Pain Management Concerns

One of the most consistent themes across most of the users was pain management. Users would mention feeling “scared” of recurrence and of “living with it for a lifetime.” They would write about the intensity of pain and the difficulty in managing pain. Many directly asked Wysa for pain relief techniques and tools, for example, “Have you got any coping skills for chronic

pain?” Some would state it as a goal for themselves, mentioning that they needed “help learning to live with chronic pain.” The want to learn pain acceptance was portrayed by a user who wrote the following:

I'm trying to be more compassionate with myself, but I have not been compassionate or patient with those parts of myself and the disorder. ... But if I can accept it anyway, I might have a chance at relief. I am listening to my body even when it is hard and even when it's not a perfect.... glamorous process. It's usually not.

Objective 2: App Engagement and Disengagement of Users With Chronic Pain

The frequency of the tools used indicated that across all tool usage by users with chronic pain within the defined time period in Wysa, users used gratitude (22%), sleep meditations (20%), anxiety management and modulation (10%), mindfulness meditations for self-compassion (10%), thought recording (7%), and conflict resolution exercises (3%) most frequently. The frequency of engagement within the app was further validated by examining the conversational flows through a network analysis mapped using cleaned and weighted data points, which depicted similar usage (shown in [Figure 2](#)).

To evaluate the intensity of engagement, the number of paths for users was calculated for both groups, along with the length of the interaction in each pathway. For paths of all lengths (any, minimal, and above the threshold), the difference in engagement for users managing chronic pain was found to be significant using the Mann-Whitney *U* test, with a large effect size, as reported in [Table 2](#).

Additionally, the analysis indicated that the points of disengagement (where users stopped engaging with the intervention) were from the same set of tools where the highest engagement was also noted. Gratitude, sleep, and thought recording were the most prominent in this list.

Figure 2. The most frequently used conversational flows for users with chronic pain (N=2194).

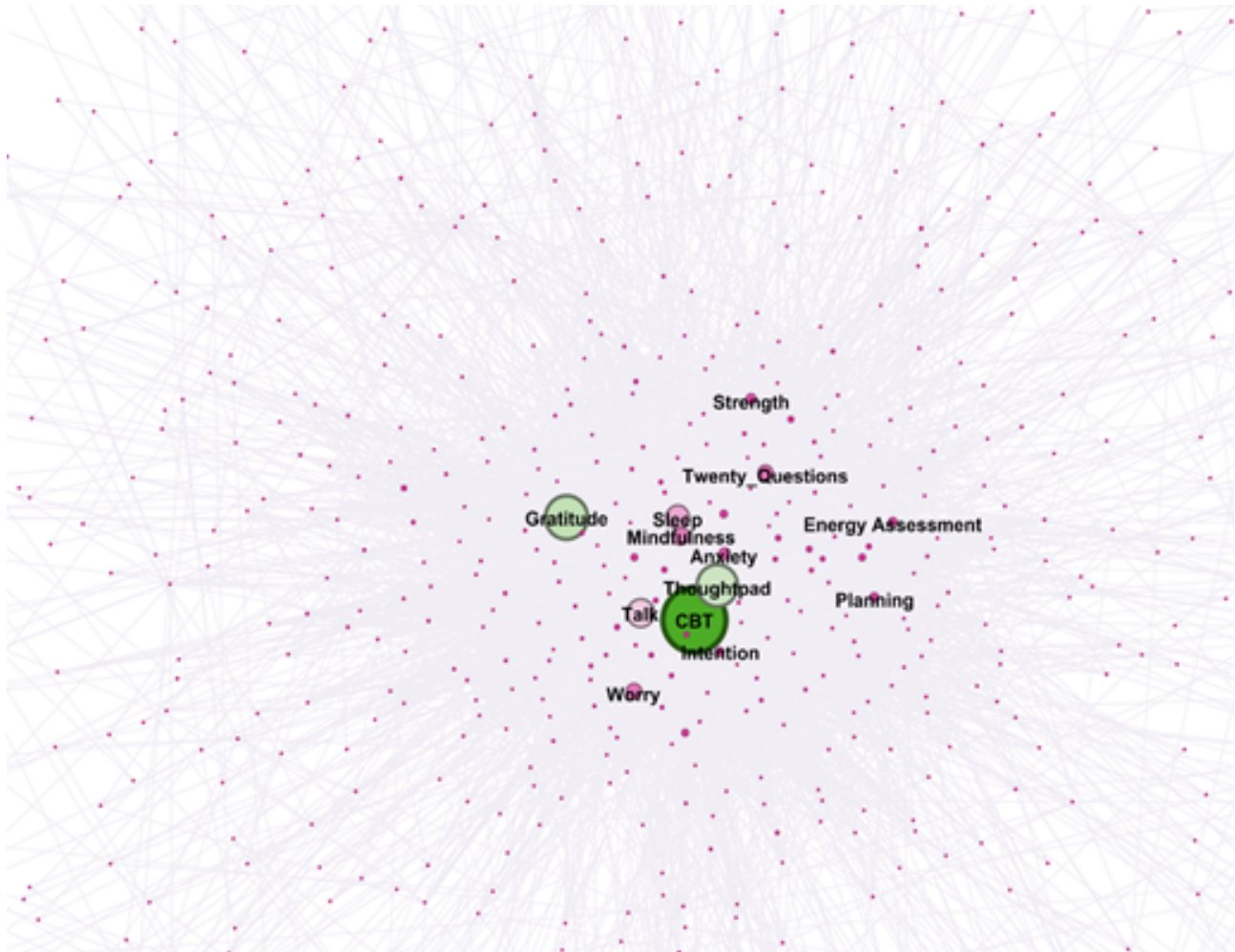


Table 2. Mann-Whitney U test for the difference in app engagement between users with chronic pain and those without chronic pain.

Variable	Users with chronic pain, n	Users without chronic pain, n	Effect size (Vargha and Delaney A)	U	P value
Paths of any length	2194	1880	0.803 (large)	3314141	<.001
Paths below the threshold for conversational units	1737	1258	0.760 (large)	1661594	<.001
Paths above the threshold for conversational units	2185	1525	0.791 (large)	2636905	<.001

Objective 3: Efficacy in Improving the Mental Health of Users With Chronic Pain

For both PHQ-9 and GAD-7 analyses, the paired *t* test indicated that there was a significant difference in the group means of the first and last assessments (PHQ-9: mean 13.22, SD 5.04 vs mean 10.03, SD 5.55; *P*<.001; GAD-7: mean 12.02, SD 4.47 vs mean 9.04, SD 5.15; *P*<.001) at the 95% confidence level, with medium effect sizes (PHQ-9: Cohen *d*=0.600; GAD-7: Cohen *d*=0.616).

Discussion

This study aimed to evaluate the perceived needs of users, the engagement of users, and the effectiveness of Wysa for mental health outcomes among users who reported chronic pain and engaged with the CBT-based digital health app for support.

The thematic analysis indicated a number of mental health concerns that chronic pain users have, including emotional difficulties and self-image-related concerns. This observational study captured the experience of these difficulties through messages. The findings are consistent with the existing literature, which highlights the association of negative emotional experiences, such as anxiety and depression, with chronic pain [28,29]. Consistent with the literature, this study highlights the bidirectional relationship between pain and health (including physical health and mental health). The emerging themes from the qualitative analysis indicate that factors, such as stress, sleep disturbances, and anxiety about health, appear to influence the physical and mental health of users with chronic pain and also affect its relationship with pain. This relationship between mental health and chronic pain highlights the importance of the treatment of mental health conditions in chronic pain [30,31].

The conversational pathways frequented and the tools used reflected the needs that clients expressed to the AI conversational agent. For instance, the users expressed their distress regarding disturbed sleep, difficulty in managing negative thoughts, difficulty in interpersonal relationships, and self-image issues. Consistent with the perceived needs and the distress that users with chronic pain reported, the tools most used and conversational pathways most frequented involved techniques and tools that address these concerns, such as exercises for cognitive reframing, thought recording, gratitude, mindful compassion, sleep and mindfulness meditations, mindful compassion meditation, and conflict resolution exercises.

The user messages also depicted the isolating nature of pain, and how chronic pain users often found themselves with limited social interaction, and scarce understanding and support from family. It is thus not surprising that one of the most common conversational flows for need fulfillment was thought recording, where the users could openly write to Wysa about their thoughts and dilemmas. Being able to offer a space for disclosure is also perhaps indicative that AI apps are able to have authentic and human-level therapeutic bonds [32], which is perhaps reinforced because of the nonjudgemental atmosphere, complete anonymity, accessibility, and constant availability that the app provides.

An analysis of the in-app engagement of chronic pain users depicted a high need for support, evidenced by the number of sessions and the length of each session, which was significantly higher compared to nonchronic pain users on the app. The qualitative analysis further elaborated on the needs of the users by capturing the immense impact of chronic pain on their lives, which extended from their physical and mental health to the socioeconomic aspects of their lives. The themes revealed that the users have concerns far beyond medical services that are generally provided. Such needs have been documented in previous studies [33-35], and the findings of this study further reinforce this.

It is also important to note that while these are the most common conversational flows and most used tools, some of these app elements and tools also represent points at which some users disengaged with Wysa. This disengagement highlights that while there is a strong need, perhaps these users who disengage do not have the necessary resources or energy levels to engage with the app, given the overwhelming number of concerns that the users with chronic pain have to deal with, as indicated in the themes. Additionally, processes, such as catastrophizing and low frustration tolerance, have been found to be involved in the cognitive processes of chronic pain patients [36]. It is possible that low frustration tolerance along with an overwhelming number of concerns caused some chronic pain users to want quick relief solutions, as often evidenced in clinical practice [37], leading them to disengage with these tools due to frustration.

The findings of this study also indicate that users not only expressed the need for learning the skills of acceptance (as indicated in the theme *pain management concerns*), but also used the tools of CBT and Mindfulness, such as cognitive reframing exercises, gratitude exercises, mindful compassion,

sleep meditations, and mindfulness meditations, as observed through path analyses. This perhaps indicates that users with chronic pain not only found these tools useful, but also continued to take these conversational paths multiple times because of the perceived usefulness. Our study further revealed significant improvements in both anxiety and depression in a subsample of users who completed the preassessment and postassessment. This evidence suggests that individuals with chronic pain can benefit from the CBT and acceptance-based tools present in Wysa. The value of these findings is highlighted when examined in light of prior literature where approaches, such as CBT, mindfulness, and acceptance-based interventions (eg, mindfulness-based stress reduction), have been found to be efficacious in the treatment of chronic pain [30,38,39] and are recommended lines of treatment.

This study has several limitations, and the findings should be interpreted in light of these limitations. This study is limited by the retrospective observational design. As such, the sample was nonrandomized, and this approach also precluded any conclusions of the causality of the effectiveness of interventions. The study was further limited because users were not required to complete assessments, limiting the study of efficacy to a small subsample. A major limitation of this study is that the data extraction keywords used were based on the guidelines of the International Association for the Study of Pain, extant literature, and the clinical experiences of the study researchers. However, despite efforts to have multiple sources for these keywords, these may not have included relevant terms or descriptions of pain. Another limitation is that the repeated measurements for efficacy, without a control group, could have raised the risk of regression to the mean [40]. The nonrandomized sampling of the design also limits the generalizability of the findings. Additionally, the third objective of the study was based on small samples, which further limits the generalizability of the findings of effectiveness, and the results should be interpreted as preliminary outcomes.

Despite the limitations, the findings of this study have important implications. This study, through its unique approach of user-led research, has highlighted the perceived needs and digital engagement patterns of users with chronic pain. Having a comprehensive understanding of the perceived needs and most frequented conversational paths of the users with chronic pain will help toward developing specific interventions, with improved product design and user experience. This could further help to study the effects of the interventions on a larger sample, with control groups (to account for effects such as regression to the mean), and to draw more generalizable conclusions. The study shows promising preliminary results for the use of AI in ameliorating mental health concerns among people with chronic pain. The findings suggest that digital interventions involving acceptance, CBT, and mindfulness-based therapies could be effective in meeting the needs, and could begin to bridge the gap between the demand for mental health support and the lack of adequate resources or personnel.

The findings of this retrospective study help in understanding the pervasive concerns and perceived needs of chronic pain users, and the patterns of their engagement and disengagement with the CBT-based AI mental health app Wysa. The results

indicated clinically meaningful and significant improvements in the anxiety and depression scores of users with chronic pain. Though this study is limited by its retrospective design, it provides promising results for filling the gap within available treatments and supporting the needs of users with chronic pain through the use of digital mental health interventions.

Conflicts of Interest

CS and MK are employees of Wysa Inc and own equity in the company. SM has no conflict of interest to report.

References

1. Goldberg DS, McGee SJ. Pain as a global public health priority. *BMC Public Health* 2011 Oct 06;11(1):770 [FREE Full text] [doi: [10.1186/1471-2458-11-770](https://doi.org/10.1186/1471-2458-11-770)] [Medline: [21978149](https://pubmed.ncbi.nlm.nih.gov/21978149/)]
2. Elzahaf RA, Tashani OA, Unsworth BA, Johnson MI. The prevalence of chronic pain with an analysis of countries with a Human Development Index less than 0.9: a systematic review without meta-analysis. *Curr Med Res Opin* 2012 Jul;28(7):1221-1229. [doi: [10.1185/03007995.2012.703132](https://doi.org/10.1185/03007995.2012.703132)] [Medline: [22697274](https://pubmed.ncbi.nlm.nih.gov/22697274/)]
3. Mills SE, Nicolson KP, Smith BH. Chronic pain: a review of its epidemiology and associated factors in population-based studies. *Br J Anaesth* 2019 Aug;123(2):e273-e283 [FREE Full text] [doi: [10.1016/j.bja.2019.03.023](https://doi.org/10.1016/j.bja.2019.03.023)] [Medline: [31079836](https://pubmed.ncbi.nlm.nih.gov/31079836/)]
4. Tunks ER, Crook J, Weir R. Epidemiology of chronic pain with psychological comorbidity: prevalence, risk, course, and prognosis. *Can J Psychiatry* 2008 Apr 01;53(4):224-234. [doi: [10.1177/070674370805300403](https://doi.org/10.1177/070674370805300403)] [Medline: [18478825](https://pubmed.ncbi.nlm.nih.gov/18478825/)]
5. Bair MJ, Robinson RL, Katon W, Kroenke K. Depression and pain comorbidity: a literature review. *Arch Intern Med* 2003 Nov 10;163(20):2433-2445. [doi: [10.1001/archinte.163.20.2433](https://doi.org/10.1001/archinte.163.20.2433)] [Medline: [14609780](https://pubmed.ncbi.nlm.nih.gov/14609780/)]
6. Brennan F, Lohman D, Gwyther L. Access to Pain Management as a Human Right. *Am J Public Health* 2019 Jan;109(1):61-65. [doi: [10.2105/ajph.2018.304743](https://doi.org/10.2105/ajph.2018.304743)]
7. Becker WC, Dorflinger L, Edmond SN, Islam L, Heapy AA, Fraenkel L. Barriers and facilitators to use of non-pharmacological treatments in chronic pain. *BMC Fam Pract* 2017 Mar 20;18(1):41 [FREE Full text] [doi: [10.1186/s12875-017-0608-2](https://doi.org/10.1186/s12875-017-0608-2)] [Medline: [28320337](https://pubmed.ncbi.nlm.nih.gov/28320337/)]
8. Phillips CJ. The Cost and Burden of Chronic Pain. *Rev Pain* 2009 Jun;3(1):2-5 [FREE Full text] [doi: [10.1177/204946370900300102](https://doi.org/10.1177/204946370900300102)] [Medline: [26526940](https://pubmed.ncbi.nlm.nih.gov/26526940/)]
9. Hewitt S, Sephton R, Yeowell G. The Effectiveness of Digital Health Interventions in the Management of Musculoskeletal Conditions: Systematic Literature Review. *J Med Internet Res* 2020 Jun 05;22(6):e15617 [FREE Full text] [doi: [10.2196/15617](https://doi.org/10.2196/15617)] [Medline: [32501277](https://pubmed.ncbi.nlm.nih.gov/32501277/)]
10. Basavarajappa C, Chand PK. Digital Platforms for Mental Health-care Delivery. *Indian J Psychol Med* 2017;39(5):703-706 [FREE Full text] [doi: [10.4103/IJPSYM.IJPSYM_209_17](https://doi.org/10.4103/IJPSYM.IJPSYM_209_17)] [Medline: [29200576](https://pubmed.ncbi.nlm.nih.gov/29200576/)]
11. de la Vega R, Miró J. mHealth: a strategic field without a solid scientific soul. a systematic review of pain-related apps. *PLoS One* 2014 Jul;9(7):e101312 [FREE Full text] [doi: [10.1371/journal.pone.0101312](https://doi.org/10.1371/journal.pone.0101312)] [Medline: [24999983](https://pubmed.ncbi.nlm.nih.gov/24999983/)]
12. Inkster B, Sarda S, Subramanian V. An Empathy-Driven, Conversational Artificial Intelligence Agent (Wysa) for Digital Mental Well-Being: Real-World Data Evaluation Mixed-Methods Study. *JMIR Mhealth Uhealth* 2018 Nov 23;6(11):e12106 [FREE Full text] [doi: [10.2196/12106](https://doi.org/10.2196/12106)] [Medline: [30470676](https://pubmed.ncbi.nlm.nih.gov/30470676/)]
13. Kroenke K, Spitzer RL, Williams JBW. The PHQ-9: validity of a brief depression severity measure. *J Gen Intern Med* 2001 Sep;16(9):606-613 [FREE Full text] [doi: [10.1046/j.1525-1497.2001.016009606.x](https://doi.org/10.1046/j.1525-1497.2001.016009606.x)] [Medline: [11556941](https://pubmed.ncbi.nlm.nih.gov/11556941/)]
14. Spitzer RL, Kroenke K, Williams JBW, Löwe B. A brief measure for assessing generalized anxiety disorder: the GAD-7. *Arch Intern Med* 2006 May 22;166(10):1092-1097. [doi: [10.1001/archinte.166.10.1092](https://doi.org/10.1001/archinte.166.10.1092)] [Medline: [16717171](https://pubmed.ncbi.nlm.nih.gov/16717171/)]
15. Revised Common Rule. The U.S. Department of Health & Human Services. URL: <https://www.hhs.gov/ohrp/regulations-and-policy/regulations/finalized-revisions-common-rule/index.html> [accessed 2022-04-10]
16. Keogh B, Stenson K. Observational Studies: Retrospective Versus Prospective. In: Hammond F, Malec J, Nick T, Buschbacher R, editors. *Handbook for Clinical Research: Design, Statistics, and Implementation*. New York, NY, USA: Demos Medical Publishing; 2014:23-26.
17. Pérez-Sola V, Roca M, Alonso J, Gabilondo A, Hernando T, Sicras-Mainar A, et al. Economic impact of treatment-resistant depression: A retrospective observational study. *J Affect Disord* 2021 Dec 01;295:578-586 [FREE Full text] [doi: [10.1016/j.jad.2021.08.036](https://doi.org/10.1016/j.jad.2021.08.036)] [Medline: [34509073](https://pubmed.ncbi.nlm.nih.gov/34509073/)]
18. Venkatesan A, Rahimi L, Kaur M, Mosunic C. Digital Cognitive Behavior Therapy Intervention for Depression and Anxiety: Retrospective Study. *JMIR Ment Health* 2020 Aug 26;7(8):e21304 [FREE Full text] [doi: [10.2196/21304](https://doi.org/10.2196/21304)] [Medline: [32845246](https://pubmed.ncbi.nlm.nih.gov/32845246/)]
19. Kunkle S, Yip M, Ξ W, Hunt J. Evaluation of an On-Demand Mental Health System for Depression Symptoms: Retrospective Observational Study. *J Med Internet Res* 2020 Jun 18;22(6):e17902 [FREE Full text] [doi: [10.2196/17902](https://doi.org/10.2196/17902)] [Medline: [32554387](https://pubmed.ncbi.nlm.nih.gov/32554387/)]
20. Douma S, Petidis K, Doumas M, Papaefthimiou P, Triantafyllou A, Kartali N, et al. Prevalence of primary hyperaldosteronism in resistant hypertension: a retrospective observational study. *The Lancet* 2008 Jun;371(9628):1921-1926. [doi: [10.1016/s0140-6736\(08\)60834-x](https://doi.org/10.1016/s0140-6736(08)60834-x)]

21. Sagayadevan V, Lee SP, Abdin E, Vaingankar J, Chen H, Chong SA, et al. Retrospective observation of mental disorders during postpartum period: Results from the Singapore mental health study. *BMC Womens Health* 2015 Dec 16;15(1):119 [FREE Full text] [doi: [10.1186/s12905-015-0279-x](https://doi.org/10.1186/s12905-015-0279-x)] [Medline: [26674202](https://pubmed.ncbi.nlm.nih.gov/26674202/)]
22. Perlman A, Vodonos Zilberg A, Bak P, Dreyfuss M, Leventer-Roberts M, Vurembrand Y, et al. Characteristics and Symptoms of App Users Seeking COVID-19-Related Digital Health Information and Remote Services: Retrospective Cohort Study. *J Med Internet Res* 2020 Oct 20;22(10):e23197 [FREE Full text] [doi: [10.2196/23197](https://doi.org/10.2196/23197)] [Medline: [32961527](https://pubmed.ncbi.nlm.nih.gov/32961527/)]
23. Zimmermann G, Venkatesan A, Rawlings K, Seahill MD. Improved Glycemic Control With a Digital Health Intervention in Adults With Type 2 Diabetes: Retrospective Study. *JMIR Diabetes* 2021 Jun 02;6(2):e28033 [FREE Full text] [doi: [10.2196/28033](https://doi.org/10.2196/28033)] [Medline: [34075880](https://pubmed.ncbi.nlm.nih.gov/34075880/)]
24. Treede R, Rief W, Barke A, Aziz Q, Bennett MI, Benoliel R, et al. A classification of chronic pain for ICD-11. *Pain* 2015 Jun;156(6):1003-1007 [FREE Full text] [doi: [10.1097/j.pain.000000000000160](https://doi.org/10.1097/j.pain.000000000000160)] [Medline: [25844555](https://pubmed.ncbi.nlm.nih.gov/25844555/)]
25. Braun V, Clarke V. Using thematic analysis in psychology. *Qualitative Research in Psychology* 2006 Jan;3(2):77-101. [doi: [10.1191/1478088706qp063oa](https://doi.org/10.1191/1478088706qp063oa)]
26. Cherven K. *Network Graph Analysis and Visualization with Gephi*. Birmingham, UK: Packt Publishing Ltd; 2013.
27. Bastian M, Heymann S, Jacomy M. *Gephi: An Open Source Software for Exploring and Manipulating Networks*. In: *Proceedings of the International AAAI Conference on Web and Social Media*. 2009 Presented at: Third International AAAI Conference on Weblogs and Social Media; May 17–20, 2009; San Jose, CA p. 361-362 URL: <https://ojs.aaai.org/index.php/ICWSM/article/view/13937>
28. Sheng J, Liu S, Wang Y, Cui R, Zhang X. The Link between Depression and Chronic Pain: Neural Mechanisms in the Brain. *Neural Plast* 2017;2017:9724371 [FREE Full text] [doi: [10.1155/2017/9724371](https://doi.org/10.1155/2017/9724371)] [Medline: [28706741](https://pubmed.ncbi.nlm.nih.gov/28706741/)]
29. Woo AK. Depression and Anxiety in Pain. *Rev Pain* 2010 Mar;4(1):8-12 [FREE Full text] [doi: [10.1177/204946371000400103](https://doi.org/10.1177/204946371000400103)] [Medline: [26527193](https://pubmed.ncbi.nlm.nih.gov/26527193/)]
30. Hooten WM. Chronic Pain and Mental Health Disorders: Shared Neural Mechanisms, Epidemiology, and Treatment. *Mayo Clin Proc* 2016 Jul;91(7):955-970. [doi: [10.1016/j.mayocp.2016.04.029](https://doi.org/10.1016/j.mayocp.2016.04.029)] [Medline: [27344405](https://pubmed.ncbi.nlm.nih.gov/27344405/)]
31. Goesling J, Lin LA, Clauw DJ. Psychiatry and Pain Management: at the Intersection of Chronic Pain and Mental Health. *Curr Psychiatry Rep* 2018 Mar 05;20(2):12. [doi: [10.1007/s11920-018-0872-4](https://doi.org/10.1007/s11920-018-0872-4)] [Medline: [29504088](https://pubmed.ncbi.nlm.nih.gov/29504088/)]
32. Bickmore T, Gruber A, Picard R. Establishing the computer-patient working alliance in automated health behavior change interventions. *Patient Educ Couns* 2005 Oct;59(1):21-30. [doi: [10.1016/j.pec.2004.09.008](https://doi.org/10.1016/j.pec.2004.09.008)] [Medline: [16198215](https://pubmed.ncbi.nlm.nih.gov/16198215/)]
33. Chou L, Cicuttini FM, Urquhart DM, Anthony SN, Sullivan K, Seneviwickrama M, et al. People with low back pain perceive needs for non-biomedical services in workplace, financial, social and household domains: a systematic review. *J Physiother* 2018 Apr;64(2):74-83 [FREE Full text] [doi: [10.1016/j.jphys.2018.02.011](https://doi.org/10.1016/j.jphys.2018.02.011)] [Medline: [29574167](https://pubmed.ncbi.nlm.nih.gov/29574167/)]
34. Dewar A, White M, Posade S, Dillon W. Using nominal group technique to assess chronic pain, patients' perceived challenges and needs in a community health region. *Health Expect* 2003 Mar;6(1):44-52 [FREE Full text] [doi: [10.1046/j.1369-6513.2003.00208.x](https://doi.org/10.1046/j.1369-6513.2003.00208.x)] [Medline: [12603627](https://pubmed.ncbi.nlm.nih.gov/12603627/)]
35. Hadi MA, Alldred DP, Briggs M, Marczewski K, Closs SJ. 'Treated as a number, not treated as a person': a qualitative exploration of the perceived barriers to effective pain management of patients with chronic pain. *BMJ Open* 2017 Jun 12;7(6):e016454 [FREE Full text] [doi: [10.1136/bmjopen-2017-016454](https://doi.org/10.1136/bmjopen-2017-016454)] [Medline: [28606909](https://pubmed.ncbi.nlm.nih.gov/28606909/)]
36. Suso-Ribera C, Jornet-Gibert M, Ribera Canudas MV, McCracken LM, Maydeu-Olivares A, Gallardo-Pujol D. There's More Than Catastrophizing in Chronic Pain: Low Frustration Tolerance and Self-Downing Also Predict Mental Health in Chronic Pain Patients. *J Clin Psychol Med Settings* 2016 Jun 19;23(2):192-206. [doi: [10.1007/s10880-016-9454-y](https://doi.org/10.1007/s10880-016-9454-y)] [Medline: [26995738](https://pubmed.ncbi.nlm.nih.gov/26995738/)]
37. Taylor D. *Managing Patients with Chronic Pain and Opioid Addiction*. Cham: Springer International Publishing; 2015.
38. Morley S, Eccleston C, Williams A. Systematic review and meta-analysis of randomized controlled trials of cognitive behaviour therapy and behaviour therapy for chronic pain in adults, excluding headache. *Pain* 1999 Mar;80(1-2):1-13. [doi: [10.1016/s0304-3959\(98\)00255-3](https://doi.org/10.1016/s0304-3959(98)00255-3)] [Medline: [10204712](https://pubmed.ncbi.nlm.nih.gov/10204712/)]
39. Veehof MM, Trompetter HR, Bohlmeijer ET, Schreurs KMG. Acceptance- and mindfulness-based interventions for the treatment of chronic pain: a meta-analytic review. *Cogn Behav Ther* 2016 Jan 28;45(1):5-31. [doi: [10.1080/16506073.2015.1098724](https://doi.org/10.1080/16506073.2015.1098724)] [Medline: [26818413](https://pubmed.ncbi.nlm.nih.gov/26818413/)]
40. Yu R, Chen L. The need to control for regression to the mean in social psychology studies. *Front Psychol* 2014 Jan 08;5:1574 [FREE Full text] [doi: [10.3389/fpsyg.2014.01574](https://doi.org/10.3389/fpsyg.2014.01574)] [Medline: [25620951](https://pubmed.ncbi.nlm.nih.gov/25620951/)]

Abbreviations

- AI:** artificial intelligence
- CBT:** cognitive behavioral therapy
- GAD-7:** Generalized Anxiety Disorder Assessment-7
- PHQ-9:** Patient Health Questionnaire-9

Edited by A Kushniruk; submitted 14.12.21; peer-reviewed by Y Wang, B Thompson; comments to author 03.03.22; revised version received 16.03.22; accepted 21.03.22; published 27.04.22.

Please cite as:

Meheli S, Sinha C, Kadaba M

Understanding People With Chronic Pain Who Use a Cognitive Behavioral Therapy–Based Artificial Intelligence Mental Health App (Wysa): Mixed Methods Retrospective Observational Study

JMIR Hum Factors 2022;9(2):e35671

URL: <https://humanfactors.jmir.org/2022/2/e35671>

doi: [10.2196/35671](https://doi.org/10.2196/35671)

PMID: [35314422](https://pubmed.ncbi.nlm.nih.gov/35314422/)

©Saha Meheli, Chaitali Sinha, Madhura Kadaba. Originally published in JMIR Human Factors (<https://humanfactors.jmir.org>), 27.04.2022. This is an open-access article distributed under the terms of the Creative Commons Attribution License (<https://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Human Factors, is properly cited. The complete bibliographic information, a link to the original publication on <https://humanfactors.jmir.org>, as well as this copyright and license information must be included.

Original Paper

Identifying Contextual Factors and Strategies for Practice Facilitation in Primary Care Quality Improvement Using an Informatics-Driven Model: Framework Development and Mixed Methods Case Study

Jiancheng Ye¹; Donna Woods¹; Jennifer Bannon¹; Lucy Bilaver¹; Gayle Kricke¹; Megan McHugh¹; Abel Kho¹; Theresa Walunas¹

Feinberg School of Medicine, Northwestern University, Chicago, IL, United States

Corresponding Author:

Jiancheng Ye
Feinberg School of Medicine, Northwestern University
633 N Saint Clair St
Chicago, IL, 60611
United States
Phone: 1 312 503 3690
Email: jjiancheng.ye@u.northwestern.edu

Related Article:

This is a corrected version. See correction statement: <https://humanfactors.jmir.org/2022/3/e40674>

Abstract

Background: The past decade has seen increasing opportunities and efforts to integrate quality improvement into health care. Practice facilitation is a proven strategy to support redesign and improvement in primary care practices that focuses on building organizational capacity for continuous improvement. Practice leadership, staff, and practice facilitators all play important roles in supporting quality improvement in primary care. However, little is known about their perspectives on the context, enablers, barriers, and strategies that impact quality improvement initiatives.

Objective: This study aimed to develop a framework to enable assessment of contextual factors, challenges, and strategies that impact practice facilitation, clinical measure performance, and the implementation of quality improvement interventions. We also illustrated the application of the framework using a real-world case study.

Methods: We developed the TITO (task, individual, technology, and organization) framework by conducting participatory stakeholder workshops and incorporating their perspectives to identify enablers and barriers to quality improvement and practice facilitation. We conducted a case study using a mixed methods approach to demonstrate the use of the framework and describe practice facilitation and factors that impact quality improvement in a primary care practice that participated in the Healthy Hearts in the Heartland study.

Results: The proposed framework was used to organize and analyze different stakeholders' perspectives and key factors based on framework domains. The case study showed that practice leaders, staff, and practice facilitators all influenced the success of the quality improvement program. However, these participants faced different challenges and used different strategies. The framework showed that barriers stemmed from patients' social determinants of health, a lack of staff and time, and unsystematic facilitation resources, while enablers included practice culture, staff buy-in, implementation of effective practice facilitation strategies, practice capacity for change, and shared complementary resources from similar, ongoing programs.

Conclusions: Our framework provided a useful and generalizable structure to guide and support assessment of future practice facilitation projects, quality improvement initiatives, and health care intervention implementation studies. The practice leader, staff, and practice facilitator all saw value in the quality improvement program and practice facilitation. Practice facilitators are key liaisons to help the quality improvement program; they help all stakeholders work toward a shared target and leverage tailored strategies. Taking advantage of resources from competing, yet complementary, programs as additional support may accelerate the effective achievement of quality improvement goals. Practice facilitation-supported quality improvement programs may be

opportunities to assist primary care practices in achieving improved quality of care through focused and targeted efforts. The case study demonstrated how our framework can support a better understanding of contextual factors for practice facilitation, which could enable well-prepared and more successful quality improvement programs for primary care practices. Combining implementation science and informatics thinking, our TITO framework may facilitate interdisciplinary research in both fields.

(*JMIR Hum Factors* 2022;9(2):e32174) doi:[10.2196/32174](https://doi.org/10.2196/32174)

KEYWORDS

quality improvement; practice facilitation; primary care; mixed-methods; practice facilitator; informatics; electronic health record; implementation science; implementation; challenge; strategy; framework; perspective

Introduction

Practice facilitation is an implementation and coaching strategy that aims to develop the capacity of primary care practices to achieve sustained quality improvement (QI) and to address gaps in the implementation of interventions [1]. There is a growing body of evidence suggesting that QI programs that use practice facilitation can produce meaningful and positive change in primary care practices [2,3], including improvements in chronic disease processes and outcome measures for diabetes, asthma [4], cardiovascular disease, and cancer [5]. In addition, practice facilitation interventions that combine audits and feedback, educational materials, and system support are more effective than interventions that use a single approach [6], and they can also lead to a more learning-focused culture, improved work environment, and greater levels of teamwork [7]. Finally, practice facilitators, individuals who are trained to provide QI coaching, can help practices engage in QI activities and develop capacity for continuous QI [1]. Given the complexity and changeability of primary care practices, understanding the context, enablers, barriers, and strategies for implementation of practice facilitation-supported QI programs may help to drive their adoption. Although previous studies have investigated the perspectives of practice leaders and practice facilitators, [8,9] few have incorporated the perspectives of practice staff, who have different roles in the practice. A framework that integrates their feedback, experiences, and strategies with implementation science, technology, and human factor elements is essential to developing effective practice facilitation strategy models [10].

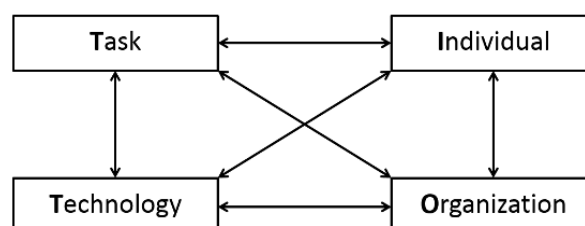
This study aims to design and develop a framework that identifies contextual factors, challenges, and strategies that impact practice facilitation, implementation of QI interventions, and clinical measure performance.

Framework Development Methods

We designed and developed the “task, individual, technology, and organization” (TITO) framework (Figure 1) by combining

the “fit between individual, task, and technology” (FITT) model [11] and the “systems engineering initiative for patient safety” (SEIPS) model [12]. The FITT model is often used to understand information technology (IT) adoption, while SEIPS is a theoretical model rooted in human-centered systems that provides a framework for understanding the structures, processes, and outcomes in health care and the relationships between these factors. The SEIPS model has been used to understand or design sociotechnical systems and has supported evaluation, planning and research activities [13]. The components in the SEIPS model include “person,” “organization,” “technologies,” “tasks,” “environment,” “process,” and “outcomes” [12]. The key stakeholders (practice leaders, practice staff, and practice facilitators), informatics researchers, and implementation scientists on the research team collaborated on the participatory workshops to develop a theory-driven framework with testable integration of the elements involved in the study. We discussed the overlaps between the two models and the unique characteristics of quality improvement research. Based on this discussion, we developed TITO by combining the FITT and SEIPS models. All the stakeholders pointed out that health IT (HIT), such as electronic health record (EHR) systems, was important for QI programs. For example, HIT includes data collection for quality measurement, patient outcome monitoring, and intervention implementation [14]. Primary care is an essential part of healthy communities. With QI programs poised to motivate clinicians to improve care quality, investment is needed to ensure that HIT used by clinicians delivers credible data on clinical quality and has the functionality necessary to inform QI efforts in addition to other purposes, such as external reporting for payment, without adding to already high burdens [15]. The research team also conducted literature reviews and multiple conversations with the research team to clarify the terminology and definitions.

Figure 1. The task, individual, technology, and organization (TITO) framework.



The FITT framework enmeshes factors related to the organization of a setting as an intrinsic part of user attributes. However, the organizational context is a critically important factor that affects both practice facilitation and intervention implementation. The “organization” dimension aids the assessment of factors related to the context in which users, tasks, and technology operate. The distinction of “organization” as a separate dimension is necessary, as this could be where key differences between different sites and settings lie. In QI, practice facilitation, or implementation science, organizational factors, such as organizational culture, readiness to engage, and capacity for change, do not fit well into either the individual, task, or technology domains. After recognizing the 4 key domains, we conducted additional literature searches, fine-tuned the domain definitions, summarized what was known about them, proposed ways to measure each domain’s use, and provided examples to increase understanding of what the domains included. Once these documents were drafted and discussed by the research team, a meeting was arranged to present each domain and discuss ways to identify questions and solicit suggestions. TITO is an informatics-driven framework based on systems thinking that can be used in various types of implementation research, such as evaluating, reporting, and synthesizing implementation studies [16].

Table 1 demonstrates the components and constructs of the TITO framework. In TITO, the “task” domain comprises the entirety of tasks and working processes (eg, data extraction and QI reports) that have to be completed by practice leaders, staff, and practice facilitators, and includes care processes, information flow, and process improvement activities. “Individual” represents key stakeholders, including practice leaders, staff, practice facilitators, and patients, as well as their physical and psychological characteristics, education, skills, knowledge, motivation, and needs. “Technology” comprises the interaction of various tools (eg, EHRs, telehealth, online training, computerized provider order entry, and medical devices) needed to accomplish the given tasks and includes electronic and digital tools, tools used by individuals to execute QI tasks, such as paper-based educational materials, and human-factor characteristics (usability, functionality, integration, and availability) [17]. Finally, “organization” includes practice culture, leadership, mission, resources, social relationships, supervisory and management style, performance evaluation, rewards and incentives, and the capacity for leading changes. Thus, the TITO framework bridges informatics and implementation science to create a testable framework for future practice facilitation projects, QI initiatives, and health care intervention implementation studies. The framework can be used to organize and analyze complex multilevel factors that impact program success.

Table 1. Components and constructs of the task, individual, technology, and organization (TITO) framework.

Domains	Examples of components and constructs
Task	General quality improvement work (data extraction and quality improvement reports), care processes, information flow, and process improvement activities
Individual	Practice leaders, practice staff, practice facilitators, physical and psychological characteristics, education, skills, knowledge, motivation, and needs
Technology	Tools (electronic health records, telehealth, online training, computerized provider order entry, and medical devices), paper-based educational materials, and human-factor characteristics (usability, functionality, integration, and availability)
Organization	Practice culture, leadership, mission, resources, social relationships, supervisory and management style, performance evaluation, rewards and incentives, and capacity for leading changes

This paper presents a case study of an application of this framework and describes context, enablers, and barriers in a primary care practice that participated in a practice facilitation–supported QI study. This case study includes perspectives from 3 key stakeholders to comprehensively examine the TITO framework, shows how each domain in the system interacts and impacts the others, and demonstrates how the framework can be used to summarize contextual factors and strategies for project success.

Case Study Implementation Methods

Healthy Hearts in the Heartland Study

The Healthy Hearts in the Heartland (H3) study aimed to examine the role of practice facilitation in improving 4 cardiovascular clinical quality measures in small primary care practices in Illinois, Indiana, and Wisconsin as part of the Agency for Healthcare Research and Quality-funded EvidenceNOW: Advancing Heart Health in Primary Care program [7]. The H3 study recruited 226 small- and medium-sized primary care practices, which were randomized

into 4 study waves that determined when they would start receiving facilitation support. Table S1 in [Multimedia Appendix 1](#) shows the characteristics of the 226 practices. Practice-tailored QI interventions were implemented over a 12-month period, followed by a 6-month sustainability phase. The 4 targeted clinical quality measures included aspirin for ischemic vascular disease, blood pressure control, cholesterol management, and smoking cessation (ABCS) [3]. The QI interventions provided by the H3 study are also shown in Table S2 of [Multimedia Appendix 1](#). Examples include reminders to order aspirin for primary prevention in appropriate patients, orders, patient instructions, patient education for home blood pressure monitoring, reminders to order prescriptions for patients with diabetes, and patient education on smoking cessation [3,18,19]. Full study details and practice characteristics have been described by Ciolino et al [3].

Practice Leaders

Practice leaders were individuals at the practice who were most familiar with the intervention and were generally physicians or

QI managers [20]. They were the champions of study implementation and assisted with the whole process of practice facilitation. Practices participating in the H3 study committed personnel time for transformation activities and data transfers for evaluation. Practice leaders monitored and managed the following activities: survey completion, engagement with H3 staff to extract data through EHR reports, troubleshooting or validating data extraction, and manual chart review.

Practice Facilitators

Practice facilitators are trained individuals who help practices develop the capacity to make meaningful changes designed to improve patients' outcomes [21]. Their work includes coaching on practice enhancement methods to facilitate system-level changes. In the H3 study, practice facilitators did the following work: conduct individual biweekly interaction with sites; train clinicians and office staff on QI methods and evidence-based tools to help implement interventions; facilitate modifications to the site's EHRs to enable systems support for ABCS measurement and monitoring; routinely engage the practice site to implement data reports to facilitate monitoring of quality performance; extract ABCS data and review data with site staff; and document intervention tracking surveys [3].

Practice Staff

Practice staff are individuals (eg, clinicians, medical assistants, and front desk staff) who work interactively with practice

facilitators to conduct the intervention activities [22]. They received structured training and coaching on clinical topics and QI strategies related to heart health. They also worked with practice facilitators to design and implement QI plans and interventions (shown in Table S2 of [Multimedia Appendix 1](#)).

Case Selection

To evaluate the TITO framework, we selected a practice from the H3 study that demonstrated an above-average improvement in performance on the ABCS measures from baseline to 12 months and follow-up performance until 18 months. This practice also performed higher than average on the implementation of QI interventions and was considered to have similar characteristics to the average practice in the H3 study across the following dimensions: (1) it had 2 to 5 clinicians, (2) it used the Epic EHR system, and (3) it was not a federally qualified health center, so it could be considered a representative practice.

Table 2 presents the characteristics of the clinical and implementation outcome measurements in this practice. The numerators (n; the number of patients meeting the ABCS criteria) and denominators (N; the total number of eligible patients at the practice for a given criterion) for each of the ABCS measures were generated from native EHR reports.

Table 2. Clinical outcome measures and implementation performance of quality improvement interventions.

Measures	Baseline	12 months	18 months
Aspirin use for at-risk individuals, n/N (%)	12/12 (100)	25/26 (96)	13/13 (100)
Blood pressure control, n/N (%)	365/415 (88)	300/339 (89)	289/338 (86)
Cholesterol management, n/N (%)	23/30 (77)	231/287 (80)	12/13 (92)
Smoking cessation, n/N (%)	127/154 (82)	188/196 (96)	1626/1661 (98)
Number of implemented interventions	19	33	34

Ethics Approval

This study was approved by the Northwestern University Institutional Review Board (STU00201720 and STU00202126). Written consent was obtained from all participants through the H3 study, which was an umbrella study.

Mixed Methods Approach

This case study applied a mixed methods approach to obtain a greater understanding of the impact of practice facilitation on QI programs, the contextual factors that enabled improved health care quality [23], the experiences of the 3 different types of stakeholder we included, and to help explain the meaning of the data and the forces that facilitated improvement in a qualitative manner [24,25]. Qualitative analyses were conducted by analyzing transcripts from semistructured interviews with practice leaders, staff, and practice facilitators to obtain their perspectives on the implementation of the QI program overall and their approaches to specific interventions. The interviews with the practice leaders and practice facilitators were conducted during the H3 study, and interviews with practice staff were conducted after the study was completed. Quantitative analyses

were based on the data from practice facilitation activities, practice surveys, and staff surveys that were collected during the H3 study.

Qualitative Data Collection and Analysis

We conducted in-depth interviews with the practice leader, the practice facilitator, and 2 practice staff members to understand their experiences and perspectives on the H3 study and to identify and organize contextual factors that impacted QI initiatives. All interviews followed a semistructured protocol (Table S3 of [Multimedia Appendix 1](#)). All interview participants had actively interacted with the H3 study.

The interviews, which were conducted by telephone, were audiotaped and transcribed. The interviews with practice staff, which were also audiotaped and transcribed, were conducted on Zoom (version 5.0) [26]. We integrated all the transcribed responses and conducted open coding and axial coding to analyze the data [27]. Two researchers (JY and JB) open-coded the interview data together to identify each instance in which participants talked about their experiences with and attitudes toward the H3 initiative. The 2 researchers then conducted axial

coding and grouped open codes that were conceptually similar. Axial coding is a qualitative research technique that involves relating data together in order to reveal codes, categories, and subcategories grounded within the participants' data [27]. For example, the category "practice culture" includes statements about a practice's organizational culture and mission; "practice facilitation" include statements describing the workflow and tasks related to practice facilitation; and "patient related barriers" includes "barriers from patients' social determinants of health and other characteristics." We resolved discrepancies and developed a consensus codebook encompassing 16 distinct codes (Table 3). The remaining transcripts were then evenly

divided between the 2 researchers and coded independently following the codebook [28].

After completing axial coding, the two researchers met and collectively identified preliminary themes. Themes that lacked representation in the data were dropped and similar themes were combined [29]. The final themes were finalized via consensus to represent the most salient perspectives of the participants. Following the proposed TITO framework, we grouped these themes into 4 categories: task, individual, technology, and organization. Under each category, we analyzed the data from 3 stakeholders: practice leader, staff, and practice facilitator.

Table 3. Healthy Hearts in the Heartland qualitative analysis codebook.

ID	Code	Definition
10—Organization		
10-1	Communication	Statements about the communication among leaders, staff, and practice facilitators.
10-2	Resource sharing	Statements about taking advantage of resources from other programs.
10-3	Practice culture	Statements about a practice's organizational culture and mission.
10-4	Capacity for change	Statements about support and mechanisms for making organizational change.
10-5	Competing priorities	Statements about competing programs or clinical tasks that impact a practice's engagement.
10-6	Lack of staff	Statements about a practice lacking personnel for completing the study.
20—Tasks		
20-1	Education and training	The instructions and support that practice facilitators provide for practice.
20-2	Practice facilitation	Statements describing the workflow and tasks related to practice facilitation.
20-3	Workload	Burdens on a practice during the quality improvement implementation.
30—Technology		
30-1	Electronic health record capacity	Functionality of the electronic health record system to support the quality improvement study practice facilitation.
30-2	Resources infrastructure	Statements about electronic or paper resources for practice facilitators and the practice.
30-3	Quality improvement report	Capacity and challenges for generating quality improvement reports.
40—Individuals		
40-1	Buy-in	Statements about practice leaders, staff, and the practice facilitator's engagement with the study.
40-2	Practice facilitator's strategy	Statements describing the practice facilitator's skills and approaches that better support practice facilitation.
40-3	Patients related barriers	Barriers from patients' social determinants of health and other characteristics.
40-4	Provider's mixed opinions.	Statements about providers' mixed opinions on the guidelines provided by the study team.

Quantitative Data Collection

Practice Facilitation Activities

During the H3 study, practice facilitators documented observations and field notes (eg, coaching activities and degree of implementation success) in standardized fields using the H3 Facilitation Activity and Intervention Tracking System (FACITS) [30]. Data in FACITS included dates of initiation and completion of relevant QI implementation outcomes, the amount of time practice facilitators spent with each practice, and activities performed during practice visits.

Practice Survey and Staff Survey

Practice surveys were completed by designated office personnel who had good insight into the clinical operations of the practice [31]. We only included records with complete answers to survey questions by the same personnel at baseline, at 12 months, and at 18 months.

The H3 study incorporated the Change Process Capability Questionnaire (CPCQ) in the practice survey. The CPCQ includes 14 items assessing the extent to which a practice has used specific QI strategies to improve cardiovascular preventive care and evaluates a practice's resiliency and capacity for change [32]. The scale has been previously validated in small practices, is reliable in measuring practice use of QI strategies, and correlates well with changes in practices and care quality

outcomes [33,34]. The CPCQ score was computed as the sum of the items, ranging from -2 (strongly disagree) to 2 (strongly agree). The overall score of the 14 items ranged from -28 to 28 [35]. Higher scores indicated greater use of QI strategies.

Results

Practice Characteristics

The selected practice had 5 clinicians (including medical doctors, nurse practitioners, and physician assistants). Before participating in the H3 study, there were no major changes at the clinic (eg, implementation of a new or different EHR system, loss of staff or managers, or moving to a new location). The practice was not in a designated medically underserved area or supporting a medically underserved population as defined by the Health Resources and Service Administration. This was a multi-specialty practice owned by a large health care system

and was neither accredited as a patient-centered medical home nor a part of an accountable care organization [36]. The practice's mission was to address chronic diseases and health disparities; the practice had participated in other QI programs, such as the WISEWOMAN (Well-Integrated Screening and Evaluation for Women Across the Nation) study [37], which shared similar goals as the H3 study, such as management and support of patients with hypertension [38].

Staff Working Status

Table 4 illustrates the number of practice members and their combined full-time equivalent (FTE) for each type of staff. FTE is the ratio of the total number of paid hours during a period divided by the number of working hours in that period (eg, one staff member working full-time and another working half time would be 2 staff and 1.5 combined FTE) [39]. FTE is often used to measure a staff member's involvement in a project or to track cost reductions in an organization [40].

Table 4. Staff working status in the practice. Some clinical staff were part-time or volunteers.

Types of staff	Value, n	Combined full-time equivalent
Clinicians, including medical doctors, doctors of osteopathic medicine, nurse practitioners, and physician assistants	4	2.8
Clinical staff providing direct patient care, including registered nurses, licensed practical nurses, medical assistants, and certified medical assistants	5	5
Office staff supporting practice operations but not involved directly with patient care, including receptionists, billing staff, and data analysts	3	3
Social workers or licensed social workers	1	1

Practice Facilitation Activities

In total, the practice facilitator conducted 39 practice facilitation activities at this practice. The total time of the activities was 805 minutes. The mean time for each activity was 57.5 (SD 26.8) minutes. Among the 39 practice facilitation activities, 11 were on site while 28 were remote. Regarding the encounter type, 20 activities were categorized as "check-in with phone or email," 16 as "QI meeting," and 3 as "other" (eg, intervention documentation or extracting data).

CPCQ Scores

The mean CPCQ score at baseline was 0 (SD 1.18); at 12 months it was 1.14 (SD 0.36), and at 18 months it was 1.86 (SD 0.36). The CPCQ results demonstrated good sustainability of improvement and capacity for leading changes at this practice. In interviews, the staff also reported that the practice had been continuing with many of the suggestions and guidance they

received from the H3 study and had continued to show improvement in the ABCS outcomes.

Participants' Feedback Summary

We analyzed and mapped the experiences of participants with the H3 study and their attitudes toward it onto the proposed TITO framework. Under each domain, we analyzed the practice survey, staff survey, and interviews. Since the practice leaders, staff, and practice facilitators had different roles in the H3 study, we examined their perspectives separately. Table 5 outlines the participants' feedback on the H3 study, based on the TITO framework, as an example of how to organize, conceptualize, and examine these contextual factors and strategies.

To demonstrate the 4 domains of the TITO framework, we will illustrate the findings from this case study in more detail to serve as an example for future studies to organize, conceptualize, and examine these contextual factors and strategies. Future studies may have different constructs under each domain.

Table 5. Summary of participant feedback on the Healthy Hearts in the Heartland study, based on the TITO (task, individual, technology, organization) framework.

Role	Task	Individual	Technology	Organization
Practice facilitator	<ul style="list-style-type: none"> Enablers: supported practice with QI^a measures and intervention implementation. Barriers: workload and complexity of the QI program tasks. 	<ul style="list-style-type: none"> Enablers: providers were willing to make changes if they found value. Barriers: providers had mixed opinions on some guidelines. 	<ul style="list-style-type: none"> Enablers: high-quality EHR^b system; inventory for personalized community resource referral list (Health Rx). Barriers: none identified. 	<ul style="list-style-type: none"> Enablers: well-prepared with rich resources and support from a large health care system. Barriers: small practice; lack of staff; competing priorities.
Practice leader	<ul style="list-style-type: none"> Enablers: scheduled monthly meeting; met with PF^c and passed on information to medical assistants and medical doctors. Barriers: workload. 	<ul style="list-style-type: none"> Enablers: interested in improving and offering better services to patients; worked well with the PF and staff. Barriers: patients' social determinants of health; patient engagement issues; time pressure 	<ul style="list-style-type: none"> Enablers: used EHR system to generate reports on QI measures. Barriers: hard copies of instructions and information were not appropriate. 	<ul style="list-style-type: none"> Enablers: practice culture facilitated positive change and improvement. Barriers: none identified.
Practice staff—nurses	<ul style="list-style-type: none"> Enablers: the program was helpful for their routine work. Barriers: some guidelines differed from those used in training at the practice. 	<ul style="list-style-type: none"> Enablers: buy-in to the intervention and coaching activities; the program provided a great deal of useful information that aligned with ongoing work; active engagement and buy-in to the QI program. Barriers: patient compliance. 	<ul style="list-style-type: none"> Enablers: satisfaction with the EHR system; regular reports kept them on track. Barriers: none identified. 	<ul style="list-style-type: none"> Enablers: the program aligned well with the practice's mission. Barriers: none identified.
Practice staff—program coordinator	<ul style="list-style-type: none"> Enablers: coordination between providers and QI programs; reaching out to patients; Spanish medical interpreter. Barriers: workload; lack of effective facilitation workflow. 	<ul style="list-style-type: none"> Enablers: the team recognized the value of the program. Barriers: patient health disparities, due to language, immigration status, or transportation issues. 	<ul style="list-style-type: none"> Enablers: support from the affiliated large health care system; satisfaction with the EHR system. Barriers: none identified. 	<ul style="list-style-type: none"> Enablers: the program aligned well with the practice's mission and ongoing work. Barriers: competing programs.

^aQI: quality improvement.

^bEHR: electronic health record.

^cPF: practice facilitator.

Tasks

Practice Facilitation

Even though the practice leader said that QI practice facilitation “was not a main priority of the practice,” the practice leader added, “it was important that we had this additional help.” The practice leader considered that the H3 study fit well with the practice's own development plan, provided needed assistance, and gave them a push to better work with resources. The leader engaged in the monthly meetings and, along with the facilitator, sat down and talked about how things were going and what could be improved. The facilitator offered suggestions and the best practice evidence that they found helpful given the current work. The leader thought that “getting an outsider's perspective on improvement is helpful.”

Intervention Implementation

The practice indicated that they wanted to implement all the H3 study interventions at the start of the study. For measures like smoking cessation, since most of the patients in this practice did not smoke, it was easy to achieve high-level performance. Cholesterol management interventions overlapped with another ongoing program in this practice, which allowed the practice to take advantage of resources. To implement the interventions, this practice's strategy was to take it one step at a time. They first worked on smoking cessation, then blood pressure control. Specifically, they focused on measures that they were struggling with. The leader said that because the practice is small, “It's easy to get distracted [by clinical work], but H3 has helped the clinic focus on quality improvement.”

Individual

Patient-Related Factors

Patients in this practice had challenges pertaining to social determinants of health [41,42]. Most were immigrants and refugees with low incomes. About 80% were primarily Spanish-speaking. The practice leader said that “patient engagement is a problem” and “transportation and work (cannot take off work) also interfere with access to care.” Given these circumstances, the leader said, “If we think we will only see the patient once, we try to take the time to emphasize what they need to do. We also try to do all the lab work during that visit.” In addition, providing hard-copy information about quitting smoking did not work well, since patients seldom read the information. The reasons included the language barrier and low interest.

Practice Facilitator

The practice facilitator for this practice had prior social work experience. The practice facilitator developed a good relationship with the practice leader and staff. The practice members trusted the practice facilitator and actively reached out with questions. If they did not see improvement, the practice facilitator remained positive and encouraging. The practice facilitator said, “If we’re not improving, maybe we’re not trying the right interventions. We’re kind of working on it together.” The practice facilitator “never forced staff members to do something they did not want.” Once the practice made improvement, the practice facilitator would “attribute the improvement to the staff.”

The practice facilitator developed the following practice facilitation working strategies: (1) after each visit, compiling a summary email that included key takeaways and next steps; (2) scheduling the time for the next meeting; (3) documenting and summarizing the meeting and what was planned for the next visit in the FACITS; (4) reviewing the previous meeting’s summary prior to the next meeting and recalling what they would be talking about; and (5) bringing additional materials or information that might be helpful.

The practice facilitator always respected the personnel in the practice, and said, “Let them lead. Don’t want them to feel like you’re not listening to them by reintroducing them to something they are already aware of” [43]. The practice facilitator formulated instructions and made sure staff knew what to do step by step. The practice facilitator also developed several effective approaches to improving engagement: (1) presenting in person and not letting the practice forget about the study because of competing priorities (the practice facilitator said, “Constant presence in a very positive way. If I ignore H3, no one else is going to pay attention”); (2) writing out definitions of clinical measurements; (3) during meetings, giving providers a paper copy of the definitions and their performance on the QI measures they were tracking so they could take notes.

The quality nurse said the practice facilitator was knowledgeable. If the practice facilitator did not know something, they would reach out to the research team and provide the information to the practice later. Even after the H3 study ended, the staff sometimes still reached out to the practice

facilitator with questions regarding some similar tasks that they had worked on before, which reinforced the sustainability of improvement. Regarding resources, the practice facilitator thought the H3 team provided an abundance of resources; however, they found it difficult to find the appropriate material when needed. The practice facilitator’s approach was to use Excel spreadsheets for audits and feedback and present the data in a way that the providers could review in a structured manner. Even so, the practice facilitator still thought that it would have been helpful to “have more of a tailored menu of ways to present the resources.”

Technology

EHR system

The EHR system used by the practice during the H3 study was Epic (version 2014, Epic Systems Corporation). The robust features of this system facilitated QI activities. The EHR vendor also helped extract data and clinical quality measures. Data from the practice physically resided in the health system’s data warehouse [44]. The EHR system was certified to meet meaningful use as defined by Health and Human Services/Office of the National Coordinator for Health Information Technology (ONC) [45]. The practice was able to incorporate clinical laboratory test results into the EHR as structured data (ie, data were recorded in discrete fields and not in text fields). The practice also had the ability to electronically share patient health information (eg, lab results, imaging reports, problem lists, and medication lists) with other providers, including hospitals, ambulatory providers, and outside labs [46].

QI Measure Report

The practice could generate reports on all four ABCS QI measures at the practice level. There was an IT service provided within the health care system that was responsible for configuring and writing quality reports from the EHRs. It also worked with the practice network, health information exchange, and hospital network to report clinical quality measures.

Organization

Infrastructure Resources

Although the practice was small, it had many resources; for example, the practice staff noted that through the WISEWOMAN program [47] “a lot of blood pressure work redesigned exam rooms through that project.” The practice was owned by a large health care system to which the practice could refer patients. It also had a very extensive patient assistance program; this program had a full-time staff member dedicated to helping patients apply for medication assistance from pharmaceutical companies.

Practice Culture

The practice was open to change and interested in improving and offering better health care services to patients. The CPCQ score in this practice increased after 12 months of practice facilitation and continued to improve during the 6-month sustainment period, which demonstrated the capacity for change and ability to maintain improvement of this practice. The leader and staff welcomed suggestions from the practice facilitator. This culture brought benefits, such as including outside

perspectives into their regular meetings and adopting best practices from other practices, as well as providing a consistent external reminder of the importance of the work. All the staff were flexible and open to new ideas and unified in the mission to address health disparities. They were always willing to support patients who faced barriers and were marginalized by the health care system. The practice leader provided strong support, and practice staff were actively engaged in the practice facilitation activities in the H3 study.

Staffing Resources

The practice leader and staff felt they had a “lack of staff.” Because it was a small clinic, they had many competing priorities.

Successful Experiences, Challenges, and Recommended Solutions

We also used the TITO framework to organize successes and challenges within the H3 study and to develop solutions to address these challenges. The results are presented in [Table 6](#).

Table 6. Summary of successful experiences, challenges, and recommended solutions.

Aspects	Successful experiences	Challenges	Recommended solutions
Task	<ul style="list-style-type: none"> Monthly meetings and discussing new strategies; everyone had a voice. Took advantage of resources from other ongoing/finished programs. Small group sessions brought back to a larger group. History of patient outreach. Informative training and education materials. Structured instructions. Interventions fit the practice’s development direction. Provided materials in the language that most patients spoke (Spanish). 	<ul style="list-style-type: none"> Providers had mixed views on some guidelines. High workload. 	<ul style="list-style-type: none"> Brainstorming sessions and discussion. Meeting over the lunch hour and catching up.
Individual	<ul style="list-style-type: none"> Practice leaders and staff were flexible and open to new strategies. Active engagement. Good relationship among practice facilitator, practice leader, and staff. Effective communication/bidirectional conversation. Practice facilitator was positive and encouraging. Quality nurse was focused. 	<ul style="list-style-type: none"> Patients’ social determinants of health and health disparities. 	<ul style="list-style-type: none"> Providing culturally competent and linguistically appropriate information about health. Incentivizing and supporting practice facilitation through improved payment models (eg, incentivize providers based on the time they work on the project and whether their progress is reasonable).
Technology	<ul style="list-style-type: none"> Well-organized electronic health record infrastructure. Inventory for personalized community resource referral list (Health Rx) enabled the practice facilitator to check what was needed. Owned by a large health system; health information technology resources were shared. 	<ul style="list-style-type: none"> Too many resources (eg, human and paper tools) for the practice facilitator. 	<ul style="list-style-type: none"> Making available resources well-organized and easy to navigate.
Organization	<ul style="list-style-type: none"> Complemented other programs. Leadership support. Focused on the mission. Understood the importance of quality improvement. High level of collaboration and teamwork. 	<ul style="list-style-type: none"> Competing programs. Limited time. Lack of staff. 	<ul style="list-style-type: none"> Complementation with resources from different programs.

Discussion

Study Overview

This study designed, developed, and piloted the TITO framework, which combined the FITT and SEIPS frameworks to understand the impact of practice facilitation on clinical measure performance and the implementation of QI interventions. We present the application of this informatics-driven framework as the analysis of a case study, describing the context, enablers, barriers, and strategies of a primary care practice that participated in a practice facilitation–supported QI program. We analyzed and compared different perspectives from 3 key stakeholders using systems thinking, which allowed for comprehensive examinations of where their perspectives aligned or diverged.

Informatics-Driven Implementation Framework

The TITO framework provides a more comprehensive description of the 4 components of QI initiatives using systems thinking (task, individual, technology, and organization). This framework could enable further development of specific measures within these domains to create a standardized template to build tailored implementation research logic models [48] and better comparisons across QI programs [49]. Because TITO was developed based on informatics perspectives and systems thinking, it may foster a common language and complement other theoretical models [50], including the Consolidated Framework for Implementation Research (CFIR) framework [51]. The tradeoff may not be significant given that qualitative results are not often considered generalizable, but rather “transferrable.” The CFIR is qualitatively different from

implementation models derived from the informatics field. Combining informatics-based thinking and implementation science models may combine the advantages of both approaches and introduce benefits for a wide variety of improvement initiatives, practice settings, and care changes. The TITO framework may provide practical and actionable guidance for different stakeholders in QI programs in primary care. For example, technologies such as EHR systems will bring benefits for tasks like QI measurement reporting. Respectful negotiations and transparent communication between practice facilitators and practice staff can foster “win-win” results. Although we applied the framework in a small primary practice and focused on QI interventions for cardiovascular care, this framework may be helpful for a wide variety of QI initiatives, practice settings, and health care systems [52,53]

Primary Care Quality Improvement

For this case study, which was an extension of the H3 study, the selected practice provided lessons that may be generalizable to a broader range of primary care practices. From the practice leader’s perspective, notable barriers included patients’ social determinants of health and a lack of staff and time, but there were also outstanding enablers, such as staff buy-in, effective practice facilitation strategies, and shared complementary resources from similar ongoing programs [54]. The practice staff thought the competing programs created a burden on their engagement and routine care. However, taking advantage of the resources from similar ongoing programs could have provided additional support, which may have helped accelerate improvement in the QI initiative. From the practice facilitator’s perspective, the key enablers were the practice’s capacity to make change and the practice culture, while notable barriers included unsystematic facilitation resources. Finally, practice leaders and staff reported benefiting from targeted assistance, such as EHR documentation guidance and connections to reporting tools, resources, and training activities. Practice facilitators, however, reported that limited engagement, busy schedules, and patient characteristics led to challenges.

Application of the Framework

Leveraging the TITO framework, we identified contextual factors and strategies for practice facilitation in primary care quality improvement in 4 domains: task, individual, technology, and organization. Overall, a successful QI program should fit well within a practice’s existing strategies and mission to enable organization-level improvement and provide appropriate assistance and resources for changes in task-level improvement [55]. In the H3 initiative, most interventions were offered based on the interests of practices in the study and what they were likely to be capable of implementing. Practice facilitation works best when the practice leader and staff actively engage with the practice facilitator, recognize the importance of the study, and agree with the implementation strategies. Effective collaboration and communication among the 3 stakeholders are essential for the successful implementation of practice facilitation and QI intervention.

For small primary care practices, the lack of staff is a major problem [56]. Our findings from this case study suggest that one way to navigate this issue is to focus efforts, implement

interventions one at a time, and use resources from other ongoing programs to complement the activity. In addition, HIT can introduce benefits with the right support [57]. An EHR system with effective reporting functionality in combination with technical support from the vendor resulted in clinical-quality-measure reports that were valuable for assessing the success of QI interventions. With a solid technology foundation, sustainable quality improvement efforts, as well as the regular collection and review of clinical measures, were readily achieved. The features illuminated by this case study may be helpful to other small primary care practices seeking to improve clinical performance.

TITO also emphasizes the individual domain; successful interventions in patient populations with health disparities may require adaptation [58]. In some instances, the primary care providers could not provide patients with appropriate care due to challenging engagement issues (eg, transportation, time, and language). This led the practice in this case study to take several actions: (1) emphasize health equity and make every effort to address any patient concerns during their clinical visits; (2) recruit volunteers who could speak the same language as the patients to reach out to individuals that had higher risks; and (3) ensure that health care providers made full use of their time during each patient visit, such as by doing all the necessary clinical care that was applicable during the visit [59]. Although patients typically are not involved in practice facilitation, it would be worthwhile to consider working with patient representatives or using a community engagement process to gather feedback on ways the QI program might best address their needs [55].

Practice facilitators are key liaisons during QI practice facilitation. They must earn trust and buy-in from the practice leader and staff from the beginning of a QI program. Developing effective communication styles and skills will help practice facilitators establish and reinforce a collaborative relationship within which they can implement and foster sustainability of the QI intervention. A commitment to collaboration with humility will go a long way in supporting practices and achieving success [55]. Practice facilitators can use motivational approaches to conduct coaching activities with clinical champions, help the practice initiate QI, facilitate the application of knowledge and QI tools to improve clinical practice, provide informational resources, and motivate practice members to engage in teamwork. Clinical staff may exhibit varying levels of acceptance of program guidelines; the practice facilitator should respect their opinions and invest in time for relationship building to understand their perspectives. In addition, the practice facilitator should use tailored strategies to manage diverse resources and ensure that materials are organized, structured, and accessible for use when needed [60] to increase the efficiency of their approach.

The TITO framework introduces “organization” as an important factor, because this could be where key differences between different sites and settings lie, especially for primary care. The presence of a practice culture with a positive attitude toward change and the absence of a disruptive level of organizational stress can be effective contributors to success. The practice should be open to change and interested in improving and

offering better services to patients, regardless of whether there are financial incentives in place. Engagement in QI initiatives is more likely to be productive when practice members actively decide to participate because the QI efforts align with their fundamental values and norms—that is, viewing targeted QI efforts as a way to provide better care to their patients—not just another revenue stream for the practice or a bothersome bureaucratic burden [60]. All the staff should be flexible, open to new ideas, and unified in their commitment to a mission to address health disparities, and practice leaders should provide strong support. With such a culture, sustainable improvement can be maintained regardless of workforce turnover.

Limitations

There are some limitations to this study. First, the interviews with the practice leader and practice facilitator were conducted during the H3 study, while interviews with staff were conducted after the initiative was completed, which may have introduced recall bias. Even so, we followed up with the practice facilitator, discussed our findings, and resolved discrepancies. Because of the timing of this investigation, we were also able to examine the sustainability of the QI initiative in this practice. Second, since this case study was focused on a single primary care practice, the study observations, results, and conclusions may not be generalizable to a wider group of practices, and the codes and categories generated from our grounded theory approach may be limited in scope. Nevertheless, this practice was selected because it had the same characteristics as most of the other

practices in the H3 study, and it could have thus provided valuable lessons and implications for practices within or outside the H3 study. Third, because of the nature of case studies, it was impossible to determine causal relationships; however, our findings could suggest hypotheses for future studies as to what contextual factors are related to success.

Conclusion

In this study, we designed and developed the TITO framework to identify contextual factors and strategies that impact practice facilitation, clinical measure performance, and implementation of QI interventions. The practice leader, staff, and practice facilitator all saw value in the QI initiative; however, they faced different challenges and used different strategies during the practice facilitation. These challenges and strategies could be clearly defined using the TITO framework. The TITO framework also supports a better understanding of the contextual factors and strategies for practice facilitation and therefore may enable better-prepared and more-successful QI programs in primary care. With the uptake of implementation science and informatics thinking, the TITO framework may facilitate interdisciplinary research in these two fields. The TITO framework will also be a useful and generalizable guideline for future practice facilitation projects, QI initiatives, and health care intervention implementation studies to organize and analyze the complex, multilevel factors that impact the success of the program.

Acknowledgments

The authors would like to thank all the practices and practice facilitators in the Healthy Hearts in the Heartland initiative, especially the participants in this case study. The authors would also like to acknowledge Mr. Richard Chagnon for the implementation of the Facilitation Activity and Intervention Tracking System to track facilitation activities. This research was funded by a grant from the Agency for Healthcare Research and Quality (1R18HS023921).

Authors' Contributions

JY conceived and designed the study, conceptualized and developed the informatics framework, and was responsible for the analyses. JY and MM led the qualitative data collection. JB contributed to the qualitative analysis. JY, JB, DW, LB, GK, MM, AK, and TW contributed to the interpretation of data. JY, JB, DW, LB, GK, MM, AK, and TW contributed to the drafting and revision of the manuscript. All the authors read and approved the final version of the manuscript.

Conflicts of Interest

Theresa Walunas receives unrelated research funding from Gilead Sciences. The other authors have no conflicts of interest to declare.

Multimedia Appendix 1
Supplemental.

[[DOCX File, 30 KB - humanfactors_v9i2e32174_app1.docx](#)]

References

1. Taylor EF, Machta RM, Meyers DS, Genevro J, Peikes DN. Enhancing the primary care team to provide redesigned care: the roles of practice facilitators and care managers. *Ann Fam Med* 2013 Jan 14;11(1):80-83 [[FREE Full text](#)] [doi: [10.1370/afm.1462](https://doi.org/10.1370/afm.1462)] [Medline: [23319510](https://pubmed.ncbi.nlm.nih.gov/23319510/)]
2. Baskerville NB, Liddy C, Hogg W. Systematic review and meta-analysis of practice facilitation within primary care settings. *Ann Fam Med* 2012 Jan 09;10(1):63-74 [[FREE Full text](#)] [doi: [10.1370/afm.1312](https://doi.org/10.1370/afm.1312)] [Medline: [22230833](https://pubmed.ncbi.nlm.nih.gov/22230833/)]

3. Ciolino JD, Jackson KL, Liss DT, Brown T, Walunas TL, Murakami L, et al. Design of healthy hearts in the heartland (H3): A practice-randomized, comparative effectiveness study. *Contemp Clin Trials* 2018 Aug;71:47-54. [doi: [10.1016/j.cct.2018.06.004](https://doi.org/10.1016/j.cct.2018.06.004)] [Medline: [29870868](https://pubmed.ncbi.nlm.nih.gov/29870868/)]
4. Grumbach K, Bainbridge E, Bodenheimer T. Facilitating improvement in primary care: the promise of practice coaching. *Issue Brief (Commonw Fund)* 2012 Jun;15:1-14. [Medline: [22712103](https://pubmed.ncbi.nlm.nih.gov/22712103/)]
5. Wang A, Pollack T, Kadziel LA, Ross SM, McHugh M, Jordan N, et al. Impact of Practice Facilitation in Primary Care on Chronic Disease Care Processes and Outcomes: a Systematic Review. *J Gen Intern Med* 2018 Nov 31;33(11):1968-1977 [FREE Full text] [doi: [10.1007/s11606-018-4581-9](https://doi.org/10.1007/s11606-018-4581-9)] [Medline: [30066117](https://pubmed.ncbi.nlm.nih.gov/30066117/)]
6. Alagoz E, Chih M, Hitchcock M, Brown R, Quanbeck A. The use of external change agents to promote quality improvement and organizational change in healthcare organizations: a systematic review. *BMC Health Serv Res* 2018 Jan 25;18(1):42 [FREE Full text] [doi: [10.1186/s12913-018-2856-9](https://doi.org/10.1186/s12913-018-2856-9)] [Medline: [29370791](https://pubmed.ncbi.nlm.nih.gov/29370791/)]
7. Agency for healthcare research and quality. *International Journal of Health Care Quality Assurance* 2001 Jul 01;14(4) [FREE Full text] [doi: [10.1108/ijhcqa.2001.06214dag.001](https://doi.org/10.1108/ijhcqa.2001.06214dag.001)]
8. McHugh M, Brown T, Liss DT, Walunas TL, Persell SD. Practice Facilitators' and Leaders' Perspectives on a Facilitated Quality Improvement Program. *Ann Fam Med* 2018 Apr 09;16(Suppl 1):S65-S71 [FREE Full text] [doi: [10.1370/afm.2197](https://doi.org/10.1370/afm.2197)] [Medline: [29632228](https://pubmed.ncbi.nlm.nih.gov/29632228/)]
9. McHugh M, Brown T, Walunas TL, Liss DT, Persell SD. Contrasting Perspectives of Practice Leaders and Practice Facilitators May Be Common in Quality Improvement Initiatives. *J Healthc Qual* 2020;42(3):e32-e38. [doi: [10.1097/JHQ.000000000000223](https://doi.org/10.1097/JHQ.000000000000223)] [Medline: [31634207](https://pubmed.ncbi.nlm.nih.gov/31634207/)]
10. Faezipour M, Ferreira S. Applying systems thinking to assess sustainability in healthcare system of systems. *Int J Syst Syst Eng* 2011;2(4):290. [doi: [10.1504/ijssse.2011.043861](https://doi.org/10.1504/ijssse.2011.043861)]
11. Ammenwerth E, Iller C, Mahler C. IT-adoption and the interaction of task, technology and individuals: a fit framework and a case study. *BMC Med Inform Decis Mak* 2006 Jan 09;6(1):3 [FREE Full text] [doi: [10.1186/1472-6947-6-3](https://doi.org/10.1186/1472-6947-6-3)] [Medline: [16401336](https://pubmed.ncbi.nlm.nih.gov/16401336/)]
12. Carayon P, Schoofs Hundt A, Karsh B, Gurses AP, Alvarado CJ, Smith M, et al. Work system design for patient safety: the SEIPS model. *Qual Saf Health Care* 2006 Dec 01;15 Suppl 1(suppl 1):i50-i58 [FREE Full text] [doi: [10.1136/qshc.2005.015842](https://doi.org/10.1136/qshc.2005.015842)] [Medline: [17142610](https://pubmed.ncbi.nlm.nih.gov/17142610/)]
13. Wooldridge AR, Carayon P, Hundt AS, Hoonakker PL. SEIPS-based process modeling in primary care. *Appl Ergon* 2017 Apr;60:240-254 [FREE Full text] [doi: [10.1016/j.apergo.2016.11.010](https://doi.org/10.1016/j.apergo.2016.11.010)] [Medline: [28166883](https://pubmed.ncbi.nlm.nih.gov/28166883/)]
14. Cohen DJ, Dorr DA, Knierim K, DuBard CA, Hemler JR, Hall JD, et al. Primary Care Practices' Abilities And Challenges In Using Electronic Health Record Data For Quality Improvement. *Health Aff (Millwood)* 2018 Apr;37(4):635-643 [FREE Full text] [doi: [10.1377/hlthaff.2017.1254](https://doi.org/10.1377/hlthaff.2017.1254)] [Medline: [29608365](https://pubmed.ncbi.nlm.nih.gov/29608365/)]
15. Ye J, Sanchez-Pinto LN. Three Data-Driven Phenotypes of Multiple Organ Dysfunction Syndrome Preserved from Early Childhood to Middle Adulthood. *AMIA Annu Symp Proc* 2020;2020:1345-1353 [FREE Full text] [Medline: [33936511](https://pubmed.ncbi.nlm.nih.gov/33936511/)]
16. Jiancheng Y. Design and development of an informatics-driven implementation research framework for primary care studies. *AMIA Annu Symp Proc* 2021;2021:1208-1214 [FREE Full text] [Medline: [35308925](https://pubmed.ncbi.nlm.nih.gov/35308925/)]
17. Ye J. The Role of Health Technology and Informatics in a Global Public Health Emergency: Practices and Implications From the COVID-19 Pandemic. *JMIR Med Inform* 2020 Jul 14;8(7):e19866 [FREE Full text] [doi: [10.2196/19866](https://doi.org/10.2196/19866)] [Medline: [32568725](https://pubmed.ncbi.nlm.nih.gov/32568725/)]
18. Greenwood DA, Litchman ML, Isaacs D, Blanchette JE, Dickinson JK, Hughes A, et al. A New Taxonomy for Technology-Enabled Diabetes Self-Management Interventions: Results of an Umbrella Review. *J Diabetes Sci Technol* 2021 Aug 11:19322968211036430. [doi: [10.1177/19322968211036430](https://doi.org/10.1177/19322968211036430)] [Medline: [34378424](https://pubmed.ncbi.nlm.nih.gov/34378424/)]
19. Ye J, Li N, Lu Y, Cheng J, Xu Y. A portable urine analyzer based on colorimetric detection. *Anal Methods* 2017;9(16):2464-2471. [doi: [10.1039/c7ay00780a](https://doi.org/10.1039/c7ay00780a)]
20. Binder CR, Absenger-Helmli I, Schilling T. The reality of transdisciplinarity: a framework-based self-reflection from science and practice leaders. *Sustain Sci* 2015 Aug 13;10(4):545-562 [FREE Full text] [doi: [10.1007/s11625-015-0328-2](https://doi.org/10.1007/s11625-015-0328-2)] [Medline: [30174728](https://pubmed.ncbi.nlm.nih.gov/30174728/)]
21. DeWalt D, Powell J, Mainwaring B. Practice coaching program manual. *Aligning Forces for Quality*. 2010. URL: http://forces4quality.org/af4q/download-document/3454/Resource-pcpmanual_1.29.14.docx [accessed 2022-05-26]
22. Ploeg J, Davies B, Edwards N, Gifford W, Miller PE. Factors influencing best-practice guideline implementation: lessons learned from administrators, nursing staff, and project leaders. *Worldviews Evid Based Nurs* 2007;4(4):210-219. [doi: [10.1111/j.1741-6787.2007.00106.x](https://doi.org/10.1111/j.1741-6787.2007.00106.x)] [Medline: [18076464](https://pubmed.ncbi.nlm.nih.gov/18076464/)]
23. Yin RK. *Case study research and applications: Design and methods*. New York, NY: Sage Publications; 2017.
24. Simons H. *Case study research in practice*. New York, NY: Sage Publications; 2009.
25. Gillham B. *Case study research methods*. London, UK: Bloomsbury Publishing; 2000.
26. Gray L, Wong-Wylie G, Rempel G, Cook K. Expanding Qualitative Research Interviewing Strategies: Zoom Video Communications. *The Qualitative Report* 2020 May 15;25(5):1292-1301. [doi: [10.46743/2160-3715/2020.4212](https://doi.org/10.46743/2160-3715/2020.4212)]
27. Ozanne JL. [Review of Basics of Qualitative Research, by Strauss A and Corbin J]. *J Mark Res* 1992 Aug;29(3):382-384. [doi: [10.2307/3172751](https://doi.org/10.2307/3172751)]

28. Kendall J. Axial coding and the grounded theory controversy. *West J Nurs Res* 1999 Dec;21(6):743-757. [doi: [10.1177/019394599902100603](https://doi.org/10.1177/019394599902100603)] [Medline: [11512211](https://pubmed.ncbi.nlm.nih.gov/11512211/)]
29. Khandkar S. Open coding. University of Calgary. URL: <https://pages.cpsc.ucalgary.ca/~saul/wiki/uploads/CPSC681/open-coding.pdf> [accessed 2022-05-26]
30. Walunas TL, Ye J, Bannon J, Wang A, Kho AN, Smith JD, et al. Does coaching matter? Examining the impact of specific practice facilitation strategies on implementation of quality improvement interventions in the Healthy Hearts in the Heartland study. *Implement Sci* 2021 Mar 31;16(1):33 [FREE Full text] [doi: [10.1186/s13012-021-01100-8](https://doi.org/10.1186/s13012-021-01100-8)] [Medline: [33789696](https://pubmed.ncbi.nlm.nih.gov/33789696/)]
31. Persell S, Liss DT, Walunas TL, Ciolino JD, Ahmad FS, Brown T, et al. Effects of 2 Forms of Practice Facilitation on Cardiovascular Prevention in Primary Care: A Practice-randomized, Comparative Effectiveness Trial. *Med Care* 2020 Apr;58(4):344-351. [doi: [10.1097/MLR.0000000000001260](https://doi.org/10.1097/MLR.0000000000001260)] [Medline: [31876643](https://pubmed.ncbi.nlm.nih.gov/31876643/)]
32. Solberg LI, Asche SE, Margolis KL, Whitebird RR. Measuring an organization's ability to manage change: the change process capability questionnaire and its use for improving depression care. *Am J Med Qual* 2008 Jun;23(3):193-200. [doi: [10.1177/1062860608314942](https://doi.org/10.1177/1062860608314942)] [Medline: [18539980](https://pubmed.ncbi.nlm.nih.gov/18539980/)]
33. Solberg LI, Brekke ML, Fazio CJ, Fowles J, Jacobsen DN, Kottke TE, et al. Lessons from experienced guideline implementers: attend to many factors and use multiple strategies. *Jt Comm J Qual Improv* 2000 Apr;26(4):171-188. [doi: [10.1016/s1070-3241\(00\)26013-6](https://doi.org/10.1016/s1070-3241(00)26013-6)] [Medline: [10749003](https://pubmed.ncbi.nlm.nih.gov/10749003/)]
34. Solberg LI, Stuck LH, Crain AL, Tillema JO, Flottemesch TJ, Whitebird RR, et al. Organizational factors and change strategies associated with medical home transformation. *Am J Med Qual* 2015 May;30(4):337-344. [doi: [10.1177/1062860614532307](https://doi.org/10.1177/1062860614532307)] [Medline: [24788251](https://pubmed.ncbi.nlm.nih.gov/24788251/)]
35. Balasubramanian BA, Marino M, Cohen DJ, Ward RL, Preston A, Springer RJ, et al. Use of Quality Improvement Strategies Among Small to Medium-Size US Primary Care Practices. *Ann Fam Med* 2018 Apr 09;16(Suppl 1):S35-S43 [FREE Full text] [doi: [10.1370/afm.2172](https://doi.org/10.1370/afm.2172)] [Medline: [29632224](https://pubmed.ncbi.nlm.nih.gov/29632224/)]
36. Jackson GL, Powers BJ, Chatterjee R, Bettger JP, Kemper AR, Hasselblad V, et al. The patient centered medical home. A systematic review. *Ann Intern Med* 2013 Feb 05;158(3):169-178 [FREE Full text] [doi: [10.7326/0003-4819-158-3-201302050-00579](https://doi.org/10.7326/0003-4819-158-3-201302050-00579)] [Medline: [24779044](https://pubmed.ncbi.nlm.nih.gov/24779044/)]
37. Will JC, Farris RP, Sanders CG, Stockmyer CK, Finkelstein EA. Health promotion interventions for disadvantaged women: overview of the WISEWOMAN projects. *J Womens Health (Larchmt)* 2004 Jun;13(5):484-502. [doi: [10.1089/1540999041281025](https://doi.org/10.1089/1540999041281025)] [Medline: [15257842](https://pubmed.ncbi.nlm.nih.gov/15257842/)]
38. Ye J, Orji I, Baldrige AS, Ojo T, Shedul G, Ugwunje E, et al. Abstract MP19: Characteristics And Patterns Of Retention In Hypertension Care In Primary Care Settings: Findings From The Hypertension Treatment In Nigeria Program. *Circulation* 2022;145. [doi: [10.1161/circ.145.suppl_1.mp19](https://doi.org/10.1161/circ.145.suppl_1.mp19)]
39. Zimmermann PG. *Nursing management secrets*. Amsterdam, Netherlands: Elsevier Health Sciences; 2002.
40. Ye J. The impact of electronic health record-integrated patient-generated health data on clinician burnout. *J Am Med Inform Assoc* 2021 Apr 23;28(5):1051-1056 [FREE Full text] [doi: [10.1093/jamia/ocab017](https://doi.org/10.1093/jamia/ocab017)] [Medline: [33822095](https://pubmed.ncbi.nlm.nih.gov/33822095/)]
41. Marmot M, Wilkinson R. *Social determinants of health*. Oxford, UK: Oxford University Press; 2005.
42. Ye J, Ren Z. Using multivariate models to examine the impact of COVID-19 pandemic and gender differences on health and health care. *MedRxiv*. Posted on September 05, 2021 . [doi: [10.1101/2021.09.02.21263055](https://doi.org/10.1101/2021.09.02.21263055)]
43. Ye J. Advancing Mental Health and Psychological Support for Health Care Workers Using Digital Technologies and Platforms. *JMIR Form Res* 2021 Jun 30;5(6):e22075 [FREE Full text] [doi: [10.2196/22075](https://doi.org/10.2196/22075)] [Medline: [34106874](https://pubmed.ncbi.nlm.nih.gov/34106874/)]
44. Ye J. Health Information System's Responses to COVID-19 Pandemic in China: A National Cross-sectional Study. *Appl Clin Inform* 2021 Mar 19;12(2):399-406 [FREE Full text] [doi: [10.1055/s-0041-1728770](https://doi.org/10.1055/s-0041-1728770)] [Medline: [34010976](https://pubmed.ncbi.nlm.nih.gov/34010976/)]
45. Ye J, Yao L, Shen J, Janarthanam R, Luo Y. Predicting mortality in critically ill patients with diabetes using machine learning and clinical notes. *BMC Med Inform Decis Mak* 2020 Dec 30;20(Suppl 11):295 [FREE Full text] [doi: [10.1186/s12911-020-01318-4](https://doi.org/10.1186/s12911-020-01318-4)] [Medline: [33380338](https://pubmed.ncbi.nlm.nih.gov/33380338/)]
46. Ye J, Wang Z, Hai J. Social Networking Service, Patient-Generated Health Data, and Population Health Informatics: National Cross-sectional Study of Patterns and Implications of Leveraging Digital Technologies to Support Mental Health and Well-being. *J Med Internet Res* 2022 Apr 29;24(4):e30898 [FREE Full text] [doi: [10.2196/30898](https://doi.org/10.2196/30898)] [Medline: [35486428](https://pubmed.ncbi.nlm.nih.gov/35486428/)]
47. Kamaraju S, DeNonomie M, Sahr N, Banerjee A, Krause K, Drew E. Promoting breast health education, Access and screening mammography for women with Cultural and financial barriers: A community engagement and outreach project in the city of Milwaukee, Wisconsin (USA). *J Clin Oncol* 2016 May 20;34(15_suppl):e18060-e18060. [doi: [10.1200/jco.2016.34.15_suppl.e18060](https://doi.org/10.1200/jco.2016.34.15_suppl.e18060)]
48. Smith JD, Li DH, Rafferty MR. The Implementation Research Logic Model: a method for planning, executing, reporting, and synthesizing implementation projects. *Implement Sci* 2020 Sep 25;15(1):84 [FREE Full text] [doi: [10.1186/s13012-020-01041-8](https://doi.org/10.1186/s13012-020-01041-8)] [Medline: [32988389](https://pubmed.ncbi.nlm.nih.gov/32988389/)]
49. Solberg LI, Kuzel A, Parchman ML, Shelley DR, Dickinson WP, Walunas TL, et al. A Taxonomy for External Support for Practice Transformation. *J Am Board Fam Med* 2021 Jan 15;34(1):32-39 [FREE Full text] [doi: [10.3122/jabfm.2021.01.200225](https://doi.org/10.3122/jabfm.2021.01.200225)] [Medline: [33452080](https://pubmed.ncbi.nlm.nih.gov/33452080/)]
50. Prochaska J, Prochaska JO, Levesque DA. A transtheoretical approach to changing organizations. *Adm Policy Ment Health* 2001 Mar;28(4):247-261. [doi: [10.1023/a:1011155212811](https://doi.org/10.1023/a:1011155212811)] [Medline: [11577653](https://pubmed.ncbi.nlm.nih.gov/11577653/)]

51. Damschroder LJ, Aron DC, Keith RE, Kirsh SR, Alexander JA, Lowery JC. Fostering implementation of health services research findings into practice: a consolidated framework for advancing implementation science. *Implement Sci* 2009 Aug 07;4(1):50 [FREE Full text] [doi: [10.1186/1748-5908-4-50](https://doi.org/10.1186/1748-5908-4-50)] [Medline: [19664226](https://pubmed.ncbi.nlm.nih.gov/19664226/)]
52. Grönqvist H, Olsson EMG, Johansson B, Held C, Sjöström J, Lindahl Norberg A, et al. Fifteen Challenges in Establishing a Multidisciplinary Research Program on eHealth Research in a University Setting: A Case Study. *J Med Internet Res* 2017 May 23;19(5):e173 [FREE Full text] [doi: [10.2196/jmir.7310](https://doi.org/10.2196/jmir.7310)] [Medline: [28536090](https://pubmed.ncbi.nlm.nih.gov/28536090/)]
53. Barnett S, Jones SC, Caton T, Iverson D, Bennett S, Robinson L. Implementing a virtual community of practice for family physician training: a mixed-methods case study. *J Med Internet Res* 2014 Mar 12;16(3):e83 [FREE Full text] [doi: [10.2196/jmir.3083](https://doi.org/10.2196/jmir.3083)] [Medline: [24622292](https://pubmed.ncbi.nlm.nih.gov/24622292/)]
54. Ye J, Ma Q. The Effects and Patterns among Mobile Health, Social Determinants, and Physical Activity: A Nationally Representative Cross-Sectional Study. *AMIA Jt Summits Transl Sci Proc* 2021;2021:653-662 [FREE Full text] [Medline: [34457181](https://pubmed.ncbi.nlm.nih.gov/34457181/)]
55. Stout S, Weeg S. The practice perspective on transformation: experience and learning from the frontlines. *Med Care* 2014 Nov;52(11 Suppl 4):S23-S25. [doi: [10.1097/MLR.000000000000239](https://doi.org/10.1097/MLR.000000000000239)] [Medline: [25310634](https://pubmed.ncbi.nlm.nih.gov/25310634/)]
56. Ye J, Zhang R, Bannon JE, Wang AA, Walunas TL, Kho AN, et al. Identifying Practice Facilitation Delays and Barriers in Primary Care Quality Improvement. *J Am Board Fam Med* 2020 Sep 28;33(5):655-664 [FREE Full text] [doi: [10.3122/jabfm.2020.05.200058](https://doi.org/10.3122/jabfm.2020.05.200058)] [Medline: [32989060](https://pubmed.ncbi.nlm.nih.gov/32989060/)]
57. Higgins TC, Crosson J, Peikes D, McNellis R, Genevro J, Meyers D. Using Health Information Technology (IT) for Primary Care Quality Improvement (QI). Agency for Healthcare Research and Quality. URL: <https://www.ahrq.gov/sites/default/files/wysiwyg/ncepqr/tools/PCMH/using-health-it-to-support-qi.pdf> [accessed 2022-06-08]
58. White-Williams C, Rossi LP, Bittner VA, Driscoll A, Durant RW, Granger BB, American Heart Association Council on Cardiovascular and Stroke Nursing, Council on Epidemiology Prevention, Council on Clinical Cardiology. Addressing Social Determinants of Health in the Care of Patients With Heart Failure: A Scientific Statement From the American Heart Association. *Circulation* 2020 Jun 02;141(22):e841-e863 [FREE Full text] [doi: [10.1161/CIR.0000000000000767](https://doi.org/10.1161/CIR.0000000000000767)] [Medline: [32349541](https://pubmed.ncbi.nlm.nih.gov/32349541/)]
59. Ye J. Pediatric Mental and Behavioral Health in the Period of Quarantine and Social Distancing With COVID-19. *JMIR Pediatr Parent* 2020 Jul 28;3(2):e19867 [FREE Full text] [doi: [10.2196/19867](https://doi.org/10.2196/19867)] [Medline: [32634105](https://pubmed.ncbi.nlm.nih.gov/32634105/)]
60. Bradley EH, Holmboe ES, Mattern JA, Roumanis SA, Radford MJ, Krumholz HM. The roles of senior management in quality improvement efforts: what are the key components? *J Healthc Manag* 2003;48(1):15-28; discussion 29. [Medline: [12592866](https://pubmed.ncbi.nlm.nih.gov/12592866/)]

Abbreviations

ABCS: aspirin for ischemic vascular disease, blood pressure control, cholesterol management, and smoking cessation

CFIR: Consolidated Framework for Implementation Research

CPCQ: Change Process Capability Questionnaire

EHR: electronic health record

FACITS: Facilitation Activity and Intervention Tracking System

FITT: fit between individual, task and technology

FTE: full-time equivalent

H3: Healthy Hearts in the Heartland

HIT: health information technology

IT: information technology

QI: quality improvement

SEIPS: systems engineering initiative for patient safety

TITO: task, individual, technology, and organization

WISEWOMAN: Well-Integrated Screening and Evaluation for Women Across the Nation

Edited by A Kushniruk; submitted 16.07.21; peer-reviewed by J Lowery; comments to author 09.08.21; revised version received 11.10.21; accepted 01.05.22; published 24.06.22.

Please cite as:

Ye J, Woods D, Bannon J, Bilaver L, Kricke G, McHugh M, Kho A, Walunas T

Identifying Contextual Factors and Strategies for Practice Facilitation in Primary Care Quality Improvement Using an Informatics-Driven Model: Framework Development and Mixed Methods Case Study

JMIR Hum Factors 2022;9(2):e32174

URL: <https://humanfactors.jmir.org/2022/2/e32174>

doi: [10.2196/32174](https://doi.org/10.2196/32174)

PMID: [35749211](https://pubmed.ncbi.nlm.nih.gov/35749211/)

©Jiancheng Ye, Donna Woods, Jennifer Bannon, Lucy Bilaver, Gayle Kricke, Megan McHugh, Abel Kho, Theresa Walunas. Originally published in JMIR Human Factors (<https://humanfactors.jmir.org>), 24.06.2022. This is an open-access article distributed under the terms of the Creative Commons Attribution License (<https://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Human Factors, is properly cited. The complete bibliographic information, a link to the original publication on <https://humanfactors.jmir.org>, as well as this copyright and license information must be included.

Original Paper

A Participatory Design Approach to Develop Visualization of Wearable Actigraphy Data for Health Care Professionals: Case Study in Qatar

Kamran Khowaja^{1,2}, PhD; Wafa Waheeda Syed³, MSc; Meghna Singh⁴, MS; Shahrads Taheri^{5,6,7}, PhD; Odette Chagoury^{5,6,7}, PhD; Dena Al-Thani¹, PhD; Michaël Aupetit³, PhD

¹Information and Computing Technology Division, College of Science and Engineering, Hamad Bin Khalifa University, Education City, Qatar

²Department of Computer Science, Shaheed Zulfikar Ali Bhutto Institute of Science and Technology, Hyderabad, Pakistan

³Social Computing, Qatar Computing Research Institute, Hamad Bin Khalifa University, Education City, Qatar

⁴Department of Computer Science and Engineering, University of Minnesota, Minneapolis, MN, United States

⁵Department of Medicine, Weill Cornell Medicine, Doha, Qatar

⁶Department of Medicine, Weill Cornell Medicine, New York, NY, United States

⁷National Obesity Treatment Center, Qatar Metabolic Institute, Doha, Qatar

Corresponding Author:

Michaël Aupetit, PhD

Social Computing, Qatar Computing Research Institute

Hamad Bin Khalifa University

Research Complex Building 1

Education City, 34110

Qatar

Phone: 974 44547150

Email: maupetit@hbku.edu.qa

Abstract

Background: Several tools have been developed for health care professionals to monitor the physical activity of their patients, but most of these tools have been considering only the needs of users in North American and European countries and applicable for only specific analytic tasks. To our knowledge, no research study has utilized the participatory design (PD) approach in the Middle East region to develop such tools, involving all the stakeholders in the product development phases, and no clear use cases have been derived from such studies that could serve future development in the field.

Objective: This study aims to develop an interactive visualization tool (ActiVis) to support local health care professionals in monitoring the physical activity of their patients measured through wearable sensors, with the overall objective of improving the health of the Qatari population.

Methods: We used PD and user-centered design methodologies to develop ActiVis, including persona development, brainwriting, and heuristic walkthrough as part of user evaluation workshops; and use cases, heuristic walkthrough, interface walkthrough, and survey as part of expert evaluation sessions.

Results: We derived and validated 6 data analysis use cases targeted at specific health care professionals from a collaborative design workshop and an expert user study. These use cases led to improving the design of the ActiVis tool to support the monitoring of patients' physical activity by nurses and family doctors. The ActiVis research prototype (RP) compared favorably with the Fitbit Dashboard, showing the importance of design tools specific to end users' needs rather than relying on repurposing existing tools designed for other types of users. The use cases we derived happen to be culturally agnostic, despite our assumption that the local Muslim and Arabic culture could impact the design of such visualization tools. At last, taking a step back, we reflect on running collaborative design sessions in a multicultural environment and oil-based economy.

Conclusions: Beyond the development of the ActiVis tool, this study can serve other visualization and human-computer interaction designers in the region to prepare their design projects and encourage health care professionals to engage with designers and engineers to improve the tools they use for supporting their daily routine. The development of the ActiVis tool for nurses, and other visualization tools specific to family doctors and clinician researchers, is still ongoing and we plan to integrate them into an operational platform for health care professionals in Qatar in the near future.

KEYWORDS

participatory design; user-centered design; visualization; health care professional; persona; brainwriting; heuristic walkthrough; use case; interface walkthrough

Introduction

According to the World Health Organization (WHO) report of 2018 [1,2], lack of physical activity is the fourth leading risk factor for mortality. Physical activity reduces the risk of coronary heart disease, stroke, hypertension, depression, type 2 diabetes, and several types of cancer. Unfortunately, physical activity across many countries is declining. In the context of Qatar, researchers at Weill Cornell Medicine - Qatar (WCM-Q) conducted a study among elementary school children between ages 7 and 12 [3]. The authors found that 42.1% of these children were either obese or overweight, and their sleep was significantly shorter than children with a healthy weight. In another study on prevalent health issues among Qatari citizens and long-term residents [4], the authors found that 83% of the population undertook little to no physical activity, and almost half of the population did not do any physical exercise. Hence, there is a need to increase the physical activity of the Qatar population to reduce the risk of related diseases as mentioned in the WHO 2018 report.

Many behavioral modification programs have been developed for more than 2 decades to reduce physical inactivity [5-8]. Nowadays, technologies allow continuous recording of individual physical activity over several days. Moreover, the use of smartphones and wearable devices (smartwatches, wristbands, etc) among children, adults, or older adults has increased in the last decade. Smartphones and wearable devices are then actively used to record, measure, and monitor body movement and activities performed by an individual using global positioning system and accelerometer installed on these devices [1,9-11]. The visualization of the recorded activity data can then show the time when an individual was the most or least active throughout the day [12] and support monitoring and exploration of such activities. We focus on the design of such visualization tools in this work.

There is a growing trend in visualization studies to explore ways to represent wearable data for self-monitoring sleep [13], analyze data by health coaches [14] or researchers [15], evaluate performance dashboard for sport [16], or evaluate time-based activity graphical representations on mobile phones [17]. Some other studies explored the best approach to visualizing the data to support behavior change [18,19] or provided a visualization dashboard to help patients understand their longitudinal health data [20]. Still, the visualization of wearable data is an active research area. A natural approach to start with is to repurpose existing visualization tools such as the Fitbit Dashboard to visualize data in a health care setting, but the actual needs of health care professionals may depart substantially from the ones of the general public self-tracking their physical activities. To the best of our knowledge, there is no visualization tool specifically designed to support the health care professionals in monitoring and analyzing physical activities of patients

through their wearable actigraph data. We also could not find a set of use cases and user roles covering such needs.

Moreover, several studies [21-23] have demonstrated the importance of the cultural, social, and local context when designing medical or health care technological solutions. Despite this view, the literature on technology acceptance mostly concentrates on highly developed North American and European countries, and little is known about health technology use and data visualization in the Arab world, including the Gulf countries [24-29] such as Qatar [3,4]. Arab countries share lots of similarities, such as cultural and religious values, language, and lifestyle [30,31], and are quite different from North American and European countries. Salgado et al [32] has highlighted that culture plays a vital role from the investigation to the design or development of new methods, theories, techniques, and systems. Hence, cultural specificities were expected when we started this project and we decided to follow a participatory design (PD) approach to collect the potentially culturally specific needs of end users.

Alabdulqader et al [33] highlighted a need to reduce the cultural gap between technology designers and users by using a PD approach. PD aims to design solutions that consider the local context and culture and has been used effectively in the health/medical domain [34-44]. PD allows researchers to involve potential users of a product or technological solution in the ideation, design, development, or appropriation of the solution [35]. Kanstrup et al [35], as a part of their review, found that workshop/group sessions/focus groups, interviews, and prototyping have been more commonly used in PD sessions of health information technology. We followed this approach in our studies.

The use of opportunistic research and sampling is commonly used in health care research as it allows researchers to use the available participants or research instruments to perform research chores [45-49]. To the best of our knowledge, there is no interactive tool that has considered the needs of local health care professionals in Qatar in their regular activities. These activities include understanding and monitoring their patients and helping/assisting them to improve their physical activity, sleep, and eventually reduce obesity. Results from a previous study [4], informal discussion with the authors [3], and an approach of opportunistic research were used as a basis to design the first prototype of an interactive tool (ActiVis) to support the mentioned needs of the local health care professionals.

This paper reports on the PD and summative evaluation of a second version of the ActiVis prototype to visualize activity data from wearable devices, which meets the needs of local health care professionals for monitoring the physical activity of their patients, to improve the physical activity of the Qatar population. We use methodologies from user-centered design

[50-58] and PD for the first time to design eHealth data visualization in Qatar.

Methods

Research Protocol and User Studies

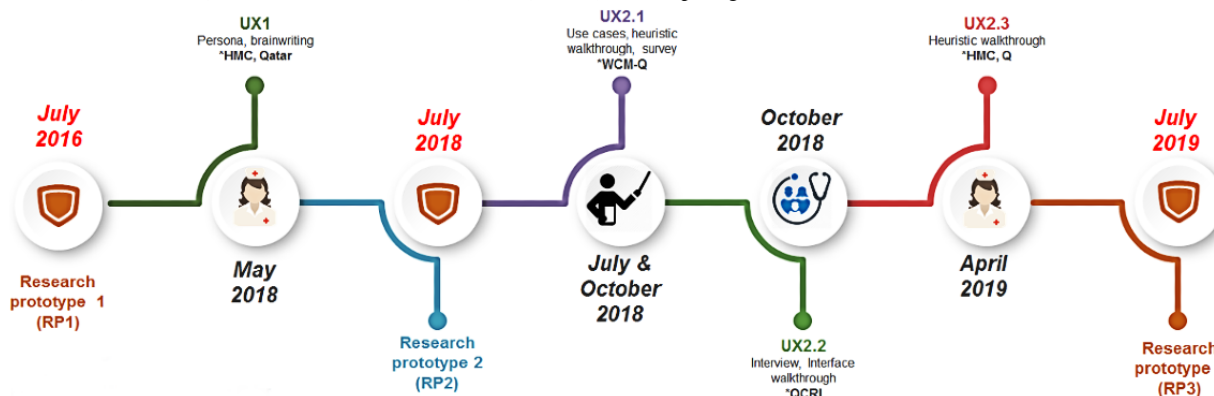
Overview

The users and their needs increased over time as studies were conducted as a part of this research and ActiVis was accordingly modified and reported in different sections of this research. The development of any technological solution is not an easy task. It requires gathering and analysis of considerable data from the

ideation to the design, development, evaluation, and deployment of the technology. It becomes even more challenging when the local context needs to be considered and incorporated into the technology. The data collection and analysis methods vary from one study to another due to various constraints such as the availability of the target users and the initial uncertainty in the direction of the project, which is refined progressively through the development cycles.

Figure 1 shows the timeline of this work, the studies conducted with their target audience, the methods used, and the venues where they took place. The RPs developed and the user studies (UX) conducted are reported in Textbox 1.

Figure 1. Research Prototypes (RP) designed and developed, and User Studies (UX) conducted throughout the project. HMC-Q: Hamad Medical Corporation - Qatar; WCM-Q: Weill Cornell Medicine - Qatar; QCRI: Qatar Computing Research Institute.



Textbox 1. Research protocols and user studies conducted.

- RP1: The first research protocol (RP) of ActiVis was developed out of a previous design expert analysis of the requirements not reported here.
- UX1: The first user study (UX) was a workshop conducted with nurses at Hamad Medical Corporation (the largest public health care provider in Qatar) to gather detailed requirements, personas, and usage scenarios, to design and develop the second RP of ActiVis (RP2) together with a set of 6 use cases targeting health care professionals.
- RP2: A total of 3 UX (UX2.1-UX2.3) were conducted to evaluate RP2 on 3 of these use cases. Each study was targeted at 1 type of user as follows:
 - UX2.1: First, an expert evaluation was conducted with clinical researchers at Weill Cornell Medicine - Qatar (WCM-Q). UX2.1 supported improving the descriptions of the use cases, determining which type of health care professional users among nurses, family doctors, and clinician researchers were the actual targets, and evaluating RP2 based on the use cases targeted at nurses and family doctors. Usability issues were also identified as a part of that study.
 - UX2.2: Then, an expert evaluation was conducted with a family doctor visiting Qatar Computing Research Institute (QCRI) to evaluate the second prototype RP2 based on use cases specific to that role as identified from UX2.1.
 - UX2.3: Lastly, a workshop was conducted with nurses from Hamad Medical Corporation. The purpose was to evaluate RP2 on the use case specifically targeted at nurses and to compare RP2 with the Fitbit Dashboard as it provided similar functionalities. The study would allow researchers to understand the differences between both dashboards from the participants' perspective and improve ActiVis based on their feedback. In this study, Fitbit was used as a comparison because it has a well-thought design [59,60] with similar functionalities required to support the user tasks, and it was the leading wearable technology in the consumer market at the time of the study [61].
- RP3: These studies (UX2.1-UX2.3) led to the design specifications for a third RP not reported here.

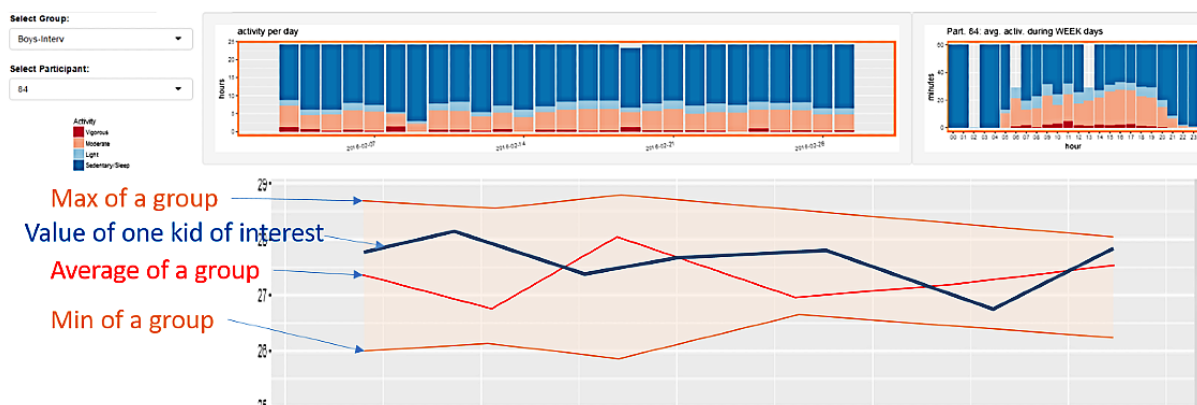
The protocol of the studies is described in the remaining subsections, while the results of each study are presented in the "Results" section.

RP1: Visual Analytic Tool for Actigraphy Sensor Data

In 2016, one of the authors (MA) started working on a visualization dashboard of wearable data for clinical decision making by health care professionals. This dashboard is aimed at supporting patients to move toward a healthier lifestyle based on their physical activity data. Figure 2 shows parts from the different screens of the initial visualization dashboard (ActiVis)

developed as an RP (RP1) based on extensive discussions with health care professionals having expertise in childhood obesity and diabetes in Qatar. The data and initial user needs to be used to design the first prototype were collected as a part of a previous research project [62,63]. The details of RP1 reported in this paper are presented in the "RP1" subsection of the "Results" section.

Figure 2. ActiVis research prototype 1.



UX1: Users' Evaluation Workshop 1 With Nurses

Overview

A first user experiment (UX1) was conducted with the nursing staff of Hamad Medical Corporation on May 2, 2018. The workshop was conducted to gather some of the potential users to generate ideas for the prototype taking the local needs into account. The objective of the workshop was to learn about nurses' perception of how visual analytics may enable them to promote lifestyle change and provide better advice to patients based on the activity data that would be collected from the patient's wearable (smartwatch). The session was focused on patients with type 2 diabetes. It included a presentation followed by a brainwriting [64] session, where nurses in groups provided their input on desired information and computer technology solutions to support patient lifestyle changes.

Findings

Our analysis of the data collected from this study led to the design of 6 use cases, and the corresponding user tasks led to the technical specifications of the visualization design that we implemented in the second prototype (RP2) of ActiVis. It is to be noted that use cases were developed from the perspective of nurses who are one of the potential users of the ActiVis tool. However, it was not clear if the description of each use case was adequate or required some improvement, and if all the use cases would need to be implemented in the ActiVis tool, justifying the needs for another set of UX (UX2.1).

Participants

A total of 45 male and female participants, which included nurses as well as nursing informatics professionals working at Hamad Medical Corporation-Qatar (HMC-Q), attended the workshop.

Study Protocol

The nursing staff working at HMC-Q were recruited through an announcement by the chairperson of the nursing department, inviting them into the workshop as shown in Figure 3 to contribute to the development of the health care solution. The participants were split into 4 groups (10-12 members in each) for the brainwriting activities. Each group was provided with a flipchart and markers in addition to in-house designed templates and gamification cards to stimulate creativity and support groups in the brainwriting process. The brainwriting process involved 4 stages:

- Stage 1: Define a "Persona"—either a nurse or a patient with diabetes. The definition must include a short biography, goals, and objectives of the persona, as well as challenges and frustrations.
- Stage 2: Describe a typical scenario, either a single encounter for the nurse or a day in the patient's life, highlighting issues and problems.
- Stage 3: Imagine the technologies that can help resolve the problems in the scenario considering the defined characteristics of the persona. Group members then vote for the best resolution.
- Stage 4: Rewrite the scenario in stage 2 including the best technology voted for in stage 3.

Figure 3. User study 1 workshop with nurses.



RP2: Research Prototype 2

The design of the second version of the ActiVis RP (RP2) was built on the use cases developed from the first user experiment

UX1. Figure 4 shows parts from the different screens of the RP2 separated by a horizontal line while the details of RP2 are presented in the “Results” section.

Figure 4. ActiVis research prototype 2.



UX2.1: Expert Evaluation 1

Overview

An expert review [65] study (UX2.1) of the second prototype (RP2) was conducted at WCQ-M. The expert review study included use cases, surveys, questionnaires, and heuristic walkthroughs. The WCM-Q group were invited for 2 studies.

UX2.1.1: Study 1

The participants went through the use cases [66,67] developed by the designers after analysis of the personas and usage scenarios from the first workshop (UX1) conducted with the nursing staff at HMC. Each participant was also asked to follow a think-aloud protocol when performing the task described in the use cases with the RP2 interface. Additional suggestions were provided toward the end of the evaluation in the survey questionnaire. The target users of the use cases were refined based on the suggestions from the participants.

UX2.1.2: Study 2

We used the heuristic walkthrough technique [68] to get participants' suggestions and improve the prototype further. The participants completed pre- and poststudy questionnaires as well. The identified usability problems were fixed before the

updated version of the prototype was further evaluated in the following user experiments (UX2.2).

Participants

The participants were working in the area of diabetes research at WCM-Q. The demographic information of the participants is shown in Table 1.

Table 1. Demographic information of the participants in UX2.1^a.

Gender	Age (years)	Position	Experience (years)	Highest degree or level of school	Competency level in computer
Male	50-59	Physician	24	Doctorate	Advanced
Female	30-39	Associate Director, Clinical Research	14	Doctorate	Intermediate
Male	30-39	Clinical trial Statistician	10	Masters	Advanced

^aUX: user study.

Study Protocol

The study protocol used was as follows:

- Participants were invited via email to be a part of the study. In the email, they were informed that the study would be conducted in-person at the campus for their ease.
- On the day of the study, the participants were briefed about the purpose of the study. The participants were informed that notes would be taken during the discussion.
- They were asked to sign a consent form before starting the study. Once signed, they were asked to complete the demographic information as part of the prestudy questionnaire (I2.1).
- The participants were asked to read through the use cases and provide suggestions on how to improve them. For each use case in the questionnaire (I2.2), the participants were asked to choose their most relevant target user, followed by a descriptive comment justifying their choice. The comments would help in making necessary changes to the use cases based on the recommendations when the use case is relevant. Additionally, they were asked 3 closed-ended questions and 1 open-ended question as described in I2.2.
- The participants were asked to evaluate the system using the heuristic walkthrough method [68]. A heuristic walkthrough is an inspection technique that combines the benefits of heuristic evaluations, cognitive walkthroughs, and usability walkthroughs [68]. It is a 2-step process. First, the participants evaluate the system based on a set of tasks and answer questions for each task based on the use cases 1, 2, and 5 from I2.2. Second, the participants identify the usability problems in the prototype and classify them using Nielsen's heuristics [69] broken down by types of usability

issues. The participants were provided a reporting template form (I2.3) to ease the process. Finally, the participants were asked to complete a poststudy questionnaire (I2.4).

Instruments Used

Overview

A total of 4 instruments were used in this study, including a prestudy questionnaire (I2.1), a use case questionnaire (I2.2), a usability problem reporting template (I2.3), and a poststudy questionnaire (I2.4). The details of each instrument and the questions included are provided in the following subsections.

I2.1: Prestudy Questionnaire

The prestudy questionnaire gathered basic information on demographic and computer skills from the participants. The questions were about gender, age, job position, university/institution/company (if a student/employed), years of experience, nationality, highest degree, and competency level of the computer.

I2.2: Use Case Questionnaire

For each use case, the participants were asked to choose the most relevant target user among 3 possible options, that is, "nurse", "clinician," and "not relevant". The participants were further asked to write a descriptive comment justifying their choice. They were also asked 3 closed-ended questions followed by 1 open-ended question. The participants had to choose the best option based on the 5-point Likert scale (1 for "strongly disagree" to 5 for "strongly agree") for each close-ended question. The open-ended question was to provide comments for the use case. The closed-ended use case questions (UCQs) are reported in Table 2.

Table 2. Closed-ended questionnaires I2.2 and I2.4.

Category and code	Text
I2.2 (Use cases)	
UCQ1 ^a	It was <i>simple</i> to use this system
UCQ2	I could effectively <i>complete</i> the tasks using this system
UCQ3	I was able to complete the tasks <i>quickly</i> using this system
I2.4 (Overall system)	
OSQ1 ^b	Overall, it was <i>easy</i> to use this system
OSQ2	It was <i>simple</i> to use this system
I2.4 (Usability)	
USBQ1 ^c	It was <i>easy</i> to learn to use this system
USBQ2	The information provided with this system was <i>clear</i> and easy to understand starting from a search query, navigating by tree keyword levels, up to getting a website description with a link to the targeted website
USBQ3	It was easy to <i>find</i> the information I needed
USBQ4	The information was effective in helping me <i>complete</i> the tasks
USBQ5	The <i>organization</i> of information on the system screens was clear
USBQ6	I <i>liked</i> using the interface of this system
I2.4 (Usefulness)	
USFQ1 ^d	This system has all the functions and capabilities I expect it to have, and
USFQ2	Overall, I am satisfied with this system performance

^aUCQ: use case question.

^bOSQ: overall system question.

^cUSB: usability question.

^dUSF: usefulness question.

I2.3: Usability Problems Reporting Template

The template provided the participants with an opportunity to report usability problems that need to be fixed in the prototype. For each usability problem, they were asked to provide a solution/recommendation from their perspective. They were also asked to add a severity rating of the problem as 0 for no problem, 1 for cosmetic, 2 for minor, 3 for major, and lastly 4 for catastrophe.

I2.4: Poststudy Questionnaire

The questionnaire contained 2 closed-ended and 1 open-ended question about the overall system usage, 6 closed-ended questions for usability, and 2 closed-ended questions on the usefulness of ActiVis. For the closed-ended questions, participants had to choose 1 option based on the 5-point Likert scale (1 for “strongly disagree” to 5 for “strongly agree”). The closed-ended questions in the 3 said categories along with the codes assigned to each question are shown in Table 2.

UX2.2: Expert Evaluation 2

Overview

A family doctor was invited to evaluate the second prototype (RP2) to realize the tasks of use cases corresponding to that role from the list refined in UX2.1.1. We followed a subset of the protocol used in UX2.1.2.

Participant

The study involved a Spanish family doctor visiting Qatar Computing Research Institute (QCRI) during October 2018, as part of his collaboration with a former investigator on this project to give feedback on QCRI’s ongoing research projects in the area of medical/health informatics. This physician was from southern Spain where a large proportion of the population are migrants from the Middle East and North Africa (ie, having Arabic origins).

Study Protocol

- The family physician was contacted through email. The participant was invited to take part in the study to share his experience and knowledge, and give feedback on the ActiVis user interface based on 3 use cases refined after UX2.1 that corresponded to the family doctor role (use cases 1, 2, and 5 were selected in Table 3). The participant acknowledged and agreed to be part of the study.
- During the study, the participant was briefed about the purpose of conducting this research and its objectives, and then introduced to the ActiVis user interface. The participant was allowed to have an informal discussion with the researcher to resolve any issues or seek any clarification before they begin the study. Written consent was also taken to be part of the evaluation.
- The participant was informed that notes would be taken throughout the study, the discussion would be

audio-recorded, the interaction during the user interface walkthrough of ActiVis would be recorded through a screen recorder application for the analysis as a backup if any point is missed while taking notes.

- The participant was informed to use a think-aloud protocol while exploring ActiVis based on the use cases. This allowed them to say out loud whatever they were thinking about how to perform a task described in each use case on ActiVis.

Table 3. Use cases (UX2.1).

Initial description resulting from the analysis of UX1 ^a by the designers and evaluated in UX2.1.1	Target user resulting from UX2.1.1
Use case 1 (Check activity level of a patient): Nurse is at her office; she gets an alert regarding patient sleep quality. Nurse accesses data of the patient; she visualizes the sleep pattern over consecutive days to check how regular it is. She detects irregular sleep time and duration with additional naps on certain days. In particular, she discovers the sleep duration is often short, and the quality of sleep is often poor. She also discovers patient activity is low to moderate.	Nurse/family doctor/clinician researcher
Use case 2 (Comparing activity between weekdays and weekends): Nurse wants to compare the average activity of the patient across weekdays and on weekends. She wants to identify irregular sleep patterns that could cause more fatigue. She discovers longer sleep duration during weekends. Also, notes that naps mostly occur around 4 PM during weekdays and around 12 PM during weekends.	Family doctor/clinician researcher
Use case 3 (Comparing 1 individual before and after intervention): Nurse compares the average activity of the patient at different periods, before and after the intervention, to assess the effectiveness of the intervention. She can see the more regular sleep pattern both during weekdays and weekends after the intervention than before it. She can also compare biometrics such as the normalized BMI and weight, between the 2 periods, and she can identify a loss of weight and decrease of BMI.	Family doctor/clinician researcher
Use case 4 (Comparing 2 individuals [siblings] over a long period): The nurse wants to compare the body metrics and sleep quality of Patient 1 aged 8 years and Patient 2 aged 10 years who are siblings, over a long period to detect a potential family lifestyle issue. The nurse compares the average activity level on weekdays and weekends, and BMI of Patient 1 and Patient 2. She observes that both follow a similar but abnormal pattern of BMI consistent with the average activity level of the corresponding periods, leading to the conclusion that it is a family lifestyle issue.	Clinician researcher
Use case 5 (Comparing an individual to a group): Nurse compares the average level of activity of the patient with the peer group of the same gender. She can see that the patient is among the overweight subgroup, although her average activity level is similar to one of the normal subgroups, leading her to conclude that the patient may have an unbalanced diet or another health issue affecting her weight.	Family doctor/clinician researcher
Use case 6 (Comparing males and females of a group before and after intervention): Nurse compares the average level of activity of 2 subgroups of different genders from a group before and after intervention to assess the effectiveness of the intervention. She can see that males increase their activity level after school during weekdays, while females increased their sleep quality, having a more stable bedtime, especially during weekends. She can also compare biometrics such as the normalized BMI and weight, between the 2 periods and she can identify a loss of weight and decrease of BMI more important for the male group.	Clinician researcher

^aUX: user study.

Instruments Used

This expert evaluation study used 2 of the instruments (I2.1 and I2.2) described in UX2.1.2.

UX2.3: Users' Evaluation Workshop 2 With Nurses

Overview

The methods used to conduct this workshop were the same as for UX1. This workshop was conducted with the nursing informatics staff at HMC-Q to evaluate the second prototype (RP2). The workshop was also conducted with the same department and at the same venue as in UX1. It was expected that some of the staff would be the same who attended the first workshop.

The purpose of conducting this workshop was to perform a summative evaluation of the latest version of the prototype and compare it with the Fitbit Dashboard, gather their qualitative feedback, and further improve the user interface.

Participants

The recruitment process of the nursing staff was the same as for UX1. A total of 45 participants, including nurses as well as nursing informatics professionals, attended the workshop.

Study Protocol

The staff of the nursing informatics department was assigned at random to 1 of 4 tables, where each table could accommodate a maximum of 10 participants. Two groups were randomly chosen and assigned to work with the Fitbit Dashboard, while the remaining 2 groups were assigned to work with the ActiVis Dashboard. All the groups were provided a laptop to explore the assigned dashboard in a web browser using temporary credentials to log-on to the dashboard. Each group was instructed to appoint 1 participant as a *group representative* who would lead the evaluation and inform them about the tasks to be performed. Each group was also instructed to nominate 1 participant as a *group secretary* who would document the entire discussion and problem found as a part of the evaluation. Each group was also given a task-driven walkthrough template.

Instruments Used

Two instruments were used in this study. These include (1) task-driven walkthrough template, and (2) heuristic evaluation of the dashboard (RP2). The details of each instrument and the questions included are provided in the following subsections. Heuristic evaluation is a usability inspection method that uses

evaluators to identify and assess the usability problems in a user interface as a part of the iterative design process. This method relies on the expertise of the domain experts to identify the usability problem in a user interface that needs to be fixed, categorize each identified problem in the heuristics, and rate its severity. The set of 10 heuristics by Nielsen [69] (Figure 5) is the most commonly used in the industry.

Figure 5. UX Check chrome extension [70] showing Nielsen's 10 heuristics [69]. UX: User study.



14.1: Task-Driven Walkthrough Template

The template contained the following 3 tasks. These 3 tasks were derived from use case 1 (Table 3) proposed after analysis by the designers of the results of the collaborative workshop with nurses (UX1) and validated as a result of UX2.1 with clinicians. Use case 1 is targeted specifically at nurses. These task numbers would be referred to in the results of UX2.3.

- Task 1: Search for the average number of steps for last week.
- Task 2: Search for average active minutes for last month.
- Task 3: Search and describe sleep patterns from May 20 to July 31, 2015.

Each group was asked to brainstorm about the steps needed to complete the task. To guide on how to come up with concrete steps, the following steps were required to complete the first task.

- Enter Patient's Name/Search in dropdown
- Navigate to Charts
- Observe the particular chart

For each step, the group was asked to answer the following questions:

- Will the user realistically be trying to do this action?
- Is the action visible?
- Will the user recognize the action as being the correct one?

Figure 6. Problems description and recommendation with UX Check [70].

The image shows a dark-themed dialog box titled "Add a heuristic". It contains the following elements:

- A dropdown menu labeled "Add a heuristic" with the selected option "Visibility of system status".
- A text area labeled "Notes".
- A text area labeled "Recommendation".
- A dropdown menu labeled "Severity" with the value "0".
- Two buttons at the bottom: "Cancel" and "Save".

Ethics Approval

The ethical approval was sought from the Qatar Biomedical Research Institute Institutional Review Board of Hamad Bin Khalifa University, Qatar, before conducting this research (QBRI-IRB 2018-019). The health care professionals as

- Will the user understand the feedback/Is the feedback appropriate?

14.2: Heuristic Evaluation of the Dashboard

For the heuristic evaluation, each group was instructed to download and add the "UX Check" [70] extension in the Google Chrome browser. This extension allows an interactive way to identify and describe the usability problems found on the web page. Opening the extension while staying on any page will show the UX Check panel on the left side of the browser as shown in Figure 5.

The extension will create the necessary regions that can be selected using a single click of the mouse. Users first need to identify any region that contains the usability problem. Clicking on the region will pop-up the dialog as shown in Figure 6. The pop-up allows users to add the heuristic problem, problem description in the form of notes, possible recommendations to fix the problem from their perspective, and lastly the severity rating. The numbers and associated description of the rating are discussed in the "Results" section. Users can save the problem for reporting or cancel their actions. The extension provides a facility for users to view all the identified problems by clicking on the "View progress" link in the pop-up shown on the left side of the web browser. They can export all the problems identified to a Microsoft Word Document by clicking on the "Export" link.

potential users were involved in all the studies as part of this research. Following the cycles of user-centered design, each study on a prototype with health care professionals provided feedback, which was used as a requirement to design an improved version as the next prototype.

Results

RP1: Visual Analytic Tool for Actigraphy Sensor Data

We developed 2 versions of the ActiVis interface. The first version (RP1) is shown in Figure 7 and was used in UX1. It was the result of the previous analysis not reported in this study. We proposed a visualization focused on 2 generic tasks: patient overview and comparison, inspired from the discussion with a previous “obesity camp” project participants, and based on the available data [62,63].

Data are body metrics (eg, BMI, weight, height) measured at regular intervals during the obesity camp, together with minute-based activity recordings from wearable accelerometers.

The interface supports an overview and comparison between the data of 2 patients, or 1 patient and a group of patients. The left panel allows selecting the patient and the body metrics features to be displayed. The right panel shows multiple line charts coding for each of the selected features through time coded on the horizontal axis. Color of the line (orange or purple) represents the selected patient or group (Figure 8). The top and bottom rows show bar charts representing the breakdown of activity levels averaged per day for the corresponding patient or group (orange or purple color of the frame; see details in Figure 9). The rightmost views show bar charts averaging the activity level per hour across the selected time window, during weekdays (first and fourth rows) and weekend days (second and third rows). The selection is done by a range selection on the central bar charts and all charts are cross-linked to focus on the same period.

Figure 7. First version (RP1) of the ActiVis tool: the left panel is used for patient and group data selection, and filtering on body metrics and activity features; the right panel shows the resulting display for overview and patient/patient and patient/group comparison.

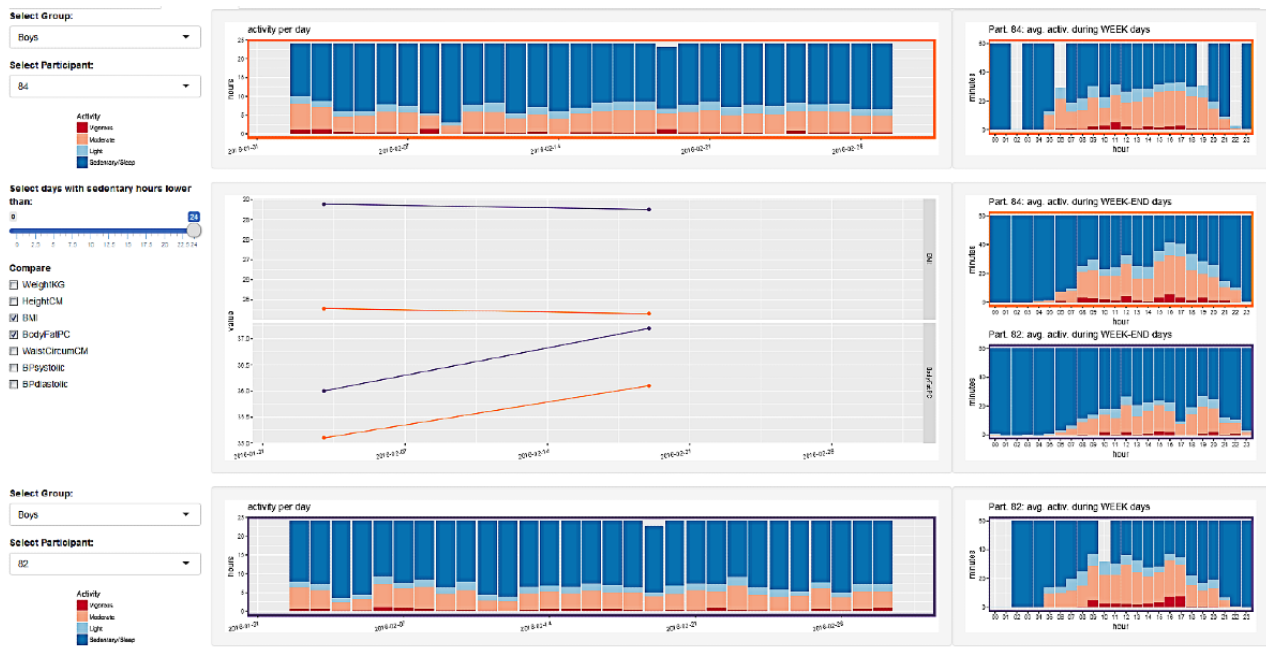


Figure 8. Details of the line chart: this chart shows the evolution of the body metric of interest (vertical axis) through time (horizontal axis) for a single patient (blue line), and a group of patients showing its minimum (orange bottom line), maximum (orange top line), and average (red line) values.

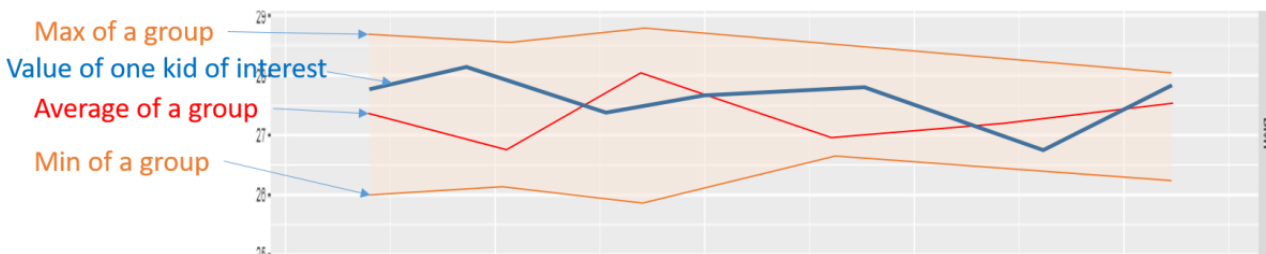
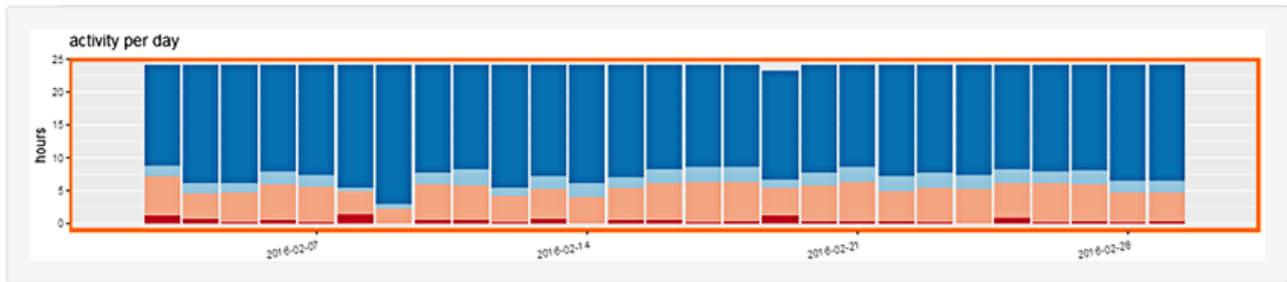


Figure 9. Details of the bar chart: Each vertical bar codes for the breakdown of activity levels per day, for sleep (dark blue), sedentary (light blue), moderate (orange), and vigorous (red) activity levels.



UX1: Users’ Evaluation Workshop 1 With Nurses

Nurses have various goals, challenges, and frustrations; however, the results showed that they are mainly concerned about patients’ awareness of their health condition and ways to monitor patients between visits. They particularly need to keep track of patients’ metrics, activity levels, and dietary habits so that they can contact the patients to guide them or remind them about what they have to do as per their activity prescription. Regarding the use of technology, some nurses raised literacy issues and others highlighted accessibility and security concerns.

Nurses highlighted that mobile health (mHealth) apps are an effective means to influence patients’ lifestyles. The most desirable functionalities are activity tracking, dietary advice, and patient education. Including a chat service to facilitate patient-nurse communication is also a viable functionality. Social networking with family and friends is crucial to encourage patients to improve their lifestyles. Interactivity features such as gamification and rewarding achievements were

identified as potential ways to motivate patients. Enabling interaction with the app and eliciting patients’ feedback facilitate tailoring contents to suit patient needs.

Outcomes of the workshop showed that recent developments in mHealth apps meet the needs and expectations of their potential users. This is consistent with the latest research findings that confirmed the popularity of mHealth apps (eg, [36]).

The analysis of the workshop usage scenarios led us to design 6 use cases reported in the left-side column of Table 3.

RP2: Research Prototype 2

Figures 10-12 show the resulting interface to support the use cases detailed in Table 3. The interface now has 3 different views to support detailed activity analysis of a patient (use case 1) in Figure 10, qualitative comparison of average activities between patients and groups of patients (use cases 2-6) in Figure 11, and quantitative analysis of the same cases in Figure 12.

Figure 10. This view supports Use Case 1: Check activity level of a patient. It shows the weekly activity of a patient broken down by day. Each row is a day, and the x-axis shows the hours from noon to noon to focus on weekly patterns of sleep (blue). The user can switch the view (top radio button) to span from midnight-to-midnight range and focus on daily activity level (reddish color). This view gives more details of each day and night, allows a side-by-side comparison, and supports the user in detecting activity patterns across several days.

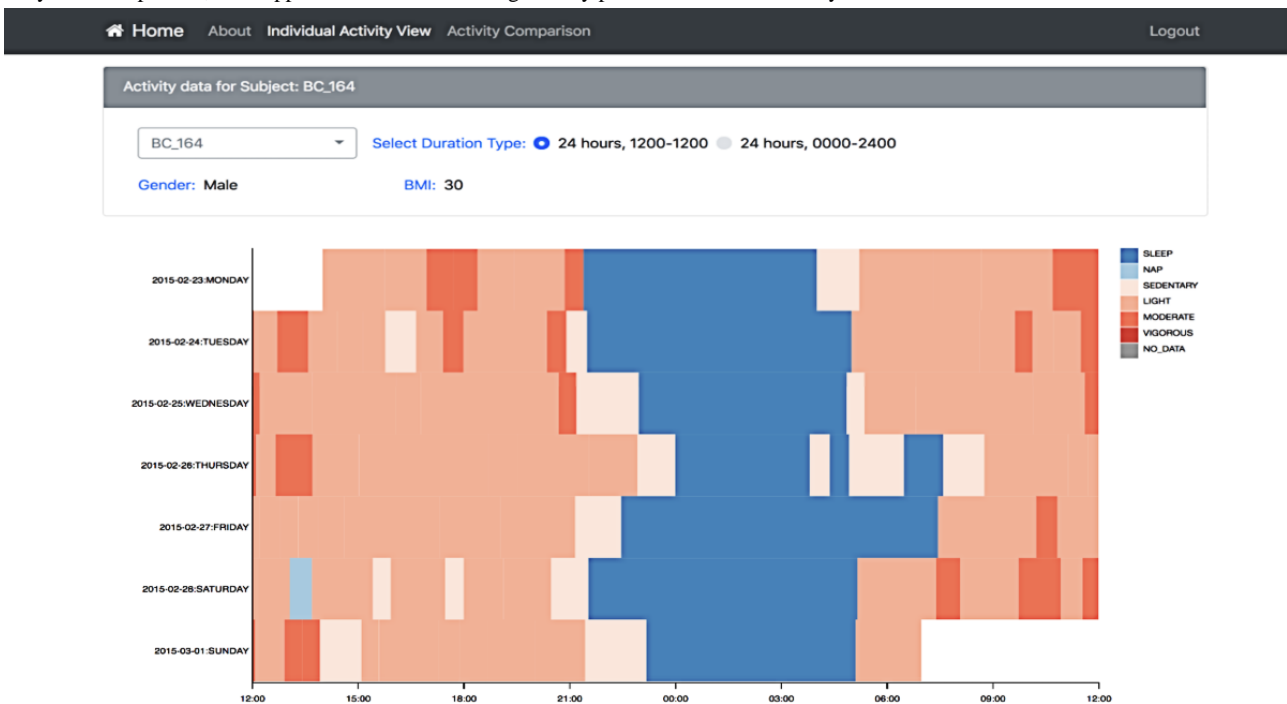


Figure 11. This view supports the qualitative pattern analysis described in Use Cases 2, 3, 4, 5, and 6. It shows a filter (top) to enable the comparison of average weekly activity between a patient or a group (left column) to another patient or a reference group (right).

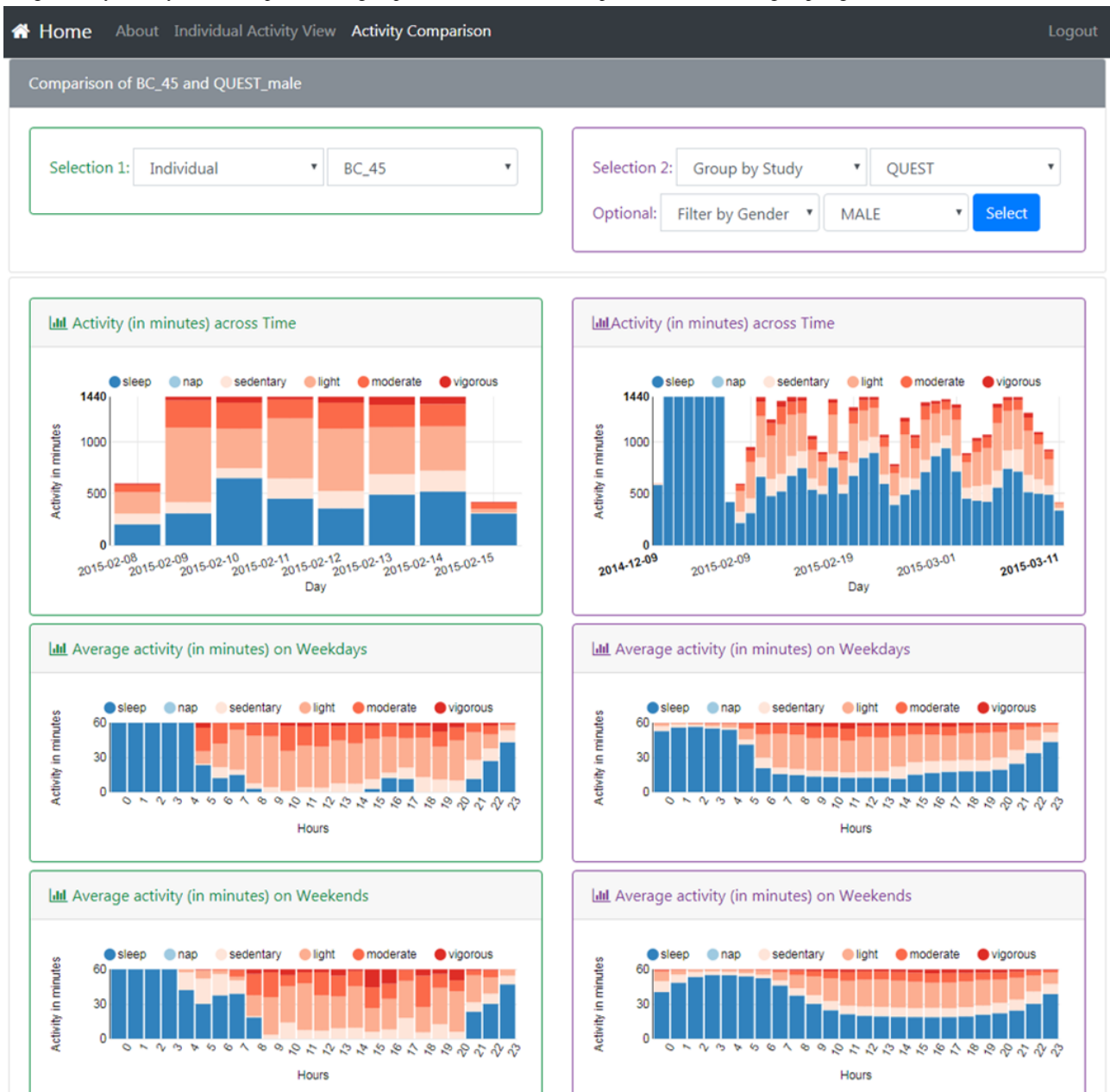


Figure 12. This view supports the quantitative pattern analysis described in Use Cases 2, 3, 4, 5, and 6. It shows the quantitative distribution of two groups of patients along different dimensions as histograms (top two rows) or combined as a color-coded scatterplot (bottom row).



UX2.1: Expert Evaluation 1

UX2.1.1: Study 1: Use Case Questionnaire

None of the participants selected an option of “Not relevant,” so all the use cases were retained and modified based on the participants’ recommendations.

Table 3 presents 6 use cases resulting from our analysis of UX1, and their reassignment to the correct target user based on the feedback of the participants in UX2.1.1.

The discussion of the results with the participants led us to further distinguish between nurses (use case 1), family doctors (use cases 1, 2, 3, and 5), and clinician researchers (all use cases) types of users. Indeed, the role of a nurse is to observe that a prescribed activity level is correctly followed by the patients to give them reminders if needed, and to notice possible anomalies to report to the doctor, both tasks falling under use case 1. The role of a family doctor is to recommend treatment to the patient.

In addition to realizing the tasks assigned to a nurse, the doctor can compare activities or biometrics of a patient between 2 periods (use case 2) to spot differences and recommend a corrective intervention to the patient. The doctors can also control the effect of their prescribed intervention by comparing activity levels and other biometrics before and after it took place (use case 3). At last, the family doctor can compare the patient with statistics derived from groups of patients with similar attributes (age, gender, BMI, or health condition; use case 5). Both nurses and doctors are focused on a single patient at a time. Finally, the clinician researcher focuses on observing trends and patterns within and between cohorts of patients (use case 6), generating knowledge that can guide the family doctors to address the health issue of a specific patient. The clinician researcher can also study more specific cases comparing them over a long period (use case 4) and in general conduct all the other tasks assigned to doctors and nurses for specific patients.

Following this refined assignment, we selected use cases 1, 2, and 5 for further summative evaluation in UX2.1.2 and UX2.2 with the family doctor, because use cases 2 and 3 involved similar tasks. We also focused strictly on use case 1 for the summative evaluation with nurses in UX2.3 as it was the only use case targeted to them.

Table 4 presents the cumulative responses related to the use case questionnaire (I2.2). The first column presents the 3 use cases used in the expert evaluation study (see Table 3 for the

full forms of mentioned use cases), the short-form of 3 questions asked for each use case is presented in the subcolumn (see I2.2 and Table 2 for the full form of each question), while the remaining columns contain the cumulative responses in terms of “strongly disagree”, “disagree”, “neutral”, “agree”, and “strongly agree.” Because 3 participants took part in the study, the maximum number of responses is less than or equal to 3. For each use case, the participants found that they were able to effectively, quickly, and efficiently complete the tasks using RP2.

Table 4. Cumulative responses of the participants for use cases 1, 2, and 5.

Use case and usability criterion	Strongly disagree	Disagree	Neutral	Agree	Strongly agree
1					
UCQ1 ^a	0	0	0	2	1
UCQ2	0	0	0	2	1
UCQ3	0	0	0	2	1
2					
UCQ1	0	0	0	2	1
UCQ2	0	0	0	1	2
UCQ3	0	0	0	1	2
5					
UCQ1	0	0	1	1	1
UCQ2	0	0	1	1	1
UCQ3	0	0	0	2	1

^aUCQ: user case question.

UX2.1.2: Study 2

Usability Problems Reporting

The descriptive comments provided by the participants as part of open-ended questions are presented in Multimedia Appendix 1. It is to be noted that minor changes were incorporated in the user interface of RP2 based on the participants' comments; therefore, no new RP was produced.

Poststudy Questionnaire

Table 5 presents the cumulative responses of the overall system, usability, and usefulness from the participant's point of view as a part of the poststudy questionnaire (see I2.4 for the questions based on the codes used in the subcolumn) using a “clustered column chart.” The format of Table 5 is similar to that of Table 4.

Table 5 shows the usefulness of the system from the participants' point of view. The participants found that the system had all the functions and capabilities they expected it to have, and they were satisfied with the performance of this system.

In terms of the overall system, Table 5 shows that the participants found that the system was easy and simple to use.

In terms of the usability of the system, Table 5 shows that the participants found that the system was easy to learn, the information provided was clear and easy to understand, the information needed was easy to find, information was effective to complete the tasks, organization of information across the screens was clear, and lastly, they liked using the interface of this system.

Table 5. Cumulative responses of the participants UX2.1.2.

Category and code	Strongly disagree	Disagree	Neutral	Agree	Strongly agree
Overall system					
OSQ1 ^a	0	0	0	1	2
OSQ2	0	0	0	2	1
Usability					
USBQ1 ^b	0	0	0	1	2
USBQ2	0	0	0	2	1
USBQ3	0	0	1	0	2
USBQ4	0	0	0	1	2
USBQ5	0	0	0	1	2
USBQ6	0	0	0	1	2
Usefulness					
USFQ1 ^c	0	0	0	2	1
USFQ2	0	0	1	1	1

^aOSQ: overall system question.

^bUSB: usability question.

^cUSF: usefulness question.

UX2.2: Expert Evaluation 2

The audio-taped RP2 interface walkthrough was analyzed. The problems identified and the recommendations provided by the participant evaluation based on the given use cases are presented in [Multimedia Appendix 2](#). Both problems and recommendations were communicated to the engineers to incorporate necessary changes in ActiVis RP2, leading to minor changes in the user interface of ActiVis RP2 used for the UX2.3.

UX2.3: Users' Evaluation Workshop 2 With Nurses

Fitbit Dashboard

Overview

[Table 6](#) shows the cumulative number of “yes” and “no” against each question for all the steps required to complete tasks 1, 2,

Table 6. Cumulative number of responses against each question for all the steps required to complete tasks using the Fitbit Dashboard.

Fitbit	Questions							
	Q1: Will the user realistically be trying to do this action?		Q2: Is the action visible?		Q3: Will user recognize the action as being the correct one?		Q4: Will the user understand the feedback/is the feedback appropriate?	
	Yes	No	Yes	No	Yes	No	Yes	No
Task 1: 6 steps	6	0	5	1	5	1	4	2
Task 2: 4 steps	4	0	4	0	4	0	4	0
Task 3: 4 steps	4	0	4	0	4	0	4	0

The results for each task are as follows:

Task 1

For all the steps in Q1, the participants were willing to perform an action. For most of the steps (5/6) in Q2 and Q3, the participants found that the action was visible, and they could

and 3 (see I4.1 for the task details) by all the groups using the Fitbit Dashboard. If the answer to any question is “yes,” then it means the group mutually agreed to the statement; however, if an answer to any question is “no,” then it shows the disagreement. In the latter case, they were instructed to add more description so that the problem can be rectified in the user interface. However, during the analysis of the filled templates returned by the groups, it was found that some of the groups also commented when their answer was “yes.” Such comments mainly reflected the minor changes recommended by the group despite an agreement to the question.

recognize that the action performed was the correct one. For 4/6 steps in Q4, the participants found that they were able to understand the feedback, or that the feedback was appropriate.

Tasks 2 and 3

For all the steps (4/4), the participants were willing to perform an action, found that the action was visible, that they recognized that the action performed was the correct one, and that the feedback given toward the end of the task was understandable or appropriate.

ActiVis Dashboard

Overview

Table 7 shows the cumulative number of “yes” and “no” against each question for all the steps required to complete tasks 1, 2, and 3 by all the groups using the ActiVis Dashboard.

The format of Table 7 is similar to that of Table 6. The results for each task are as described in the following sections.

Table 7. Cumulative number of responses against each question for all the steps required to complete tasks using the ActiVis Dashboard.

ActiVis	Questions							
	Q1: Will the user realistically be trying to do this action?		Q2: Is the action visible?		Q3: Will user recognize the action as being the correct one?		Q4: Will the user understand the feedback/Is the feedback appropriate?	
	Yes	No	Yes	No	Yes	No	Yes	No
Task 1: 6 steps	5	1	6	0	4	2	3	3
Task 2: 6 steps	5	1	3	3	4	2	4	2
Task 3: 6 steps	4	2	5	1	5	1	6	0

Task 1

For most of the steps (5/6) in Q1, the participants were willing to perform an action, for all the steps (6/6) in Q2, the participants found that the action was visible. For 4/6 steps in Q3, the participants were able to recognize that the action performed was the correct one. However, for 3/6 steps in Q4, the participants had mixed opinions; for half of the steps, they found that they were either unable to understand the feedback, or that the feedback was inappropriate, while for the remaining steps, they found that they were able to understand the feedback, or that the feedback was appropriate.

Task 2

For most of the steps (3/4) in Q1, the participants were willing to perform an action; however, for 3/5 steps in Q2, the participants found that the action was not visible. For 2/4 steps in Q3, the participants had mixed opinions. For half of the steps, some participants found that they were able to recognize the action performed, while the other participants found that they were unable to recognize the action performed. Similarly, a mixed opinion was also found for Q4 (2/4 steps). For half of the steps, some participants found that they were able to understand the feedback given toward the end of the task, while the other participants found that they were unable to understand the feedback given at the end of the task.

Task 3

For 3/5 steps in Q1, the participants were willing to perform an action, for 4/5 steps in Q2 and Q3 each, the participants found that the action was visible and that they recognized that action performed was the correct one. For all the steps, the participants found that the feedback given after the task was performed was understandable or appropriate.

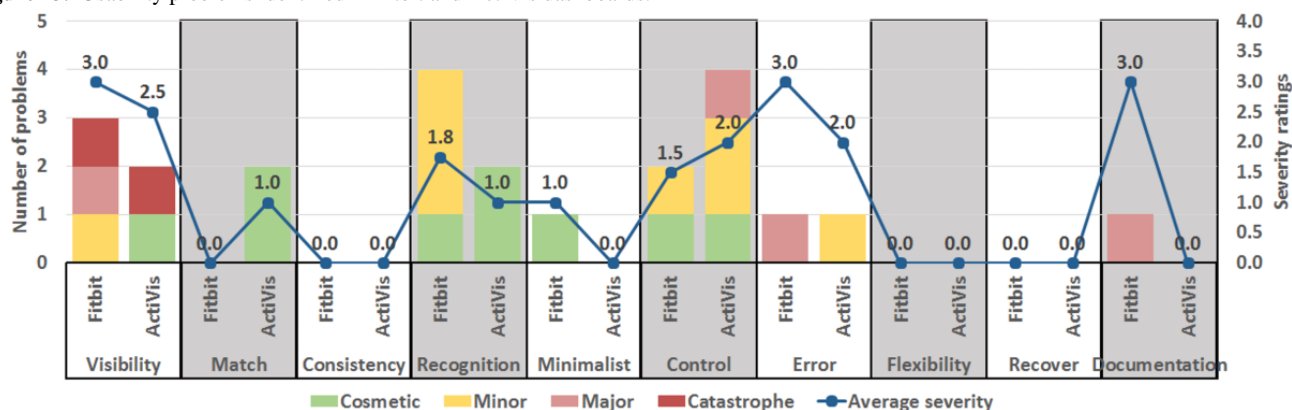
Heuristic Evaluation of the Interfaces

Figure 13 shows the number of usability problems found and the average severity ratings of the identified problems in the Fitbit Dashboard and the ActiVis Dashboard, respectively, using Nielsen’s 10 heuristics. The “stacked columns” represent the “number of usability problems” (left vertical scale), whereas the “line with markers” represents the “average severity rating of the identified problems” (right vertical scale). Each stack column shows the number of usability problems found based on the 4 severity ratings, that is, cosmetic, minor, major, and catastrophic. The axis on the left-hand side is known as the primary axis and it is related to the “stacked columns,” whereas the axis on the right-hand side is known as the secondary axis and is related to the “line with markers.”

A total of 11 usability problems were identified in each of the 2 dashboards (ie, Fitbit and ActiVis). The analysis of the results in terms of the number of usability problems found in Fitbit shows that the recognition heuristic (n=4) was the more commonly broken heuristic, followed by the visibility and control heuristics (n=2 each). Similarly, the analysis of the results in terms of the number of usability problems found in ActiVis shows that the control heuristic (n=4) was the more commonly broken heuristic, followed by the visibility, match, and recognition heuristics (n=2 each).

The analysis of the results in terms of the average severity rating shows that the majority of problems identified are minor.

The number of usability problems identified and their severity rating provided by the participants for the Fitbit Dashboard and the ActiVis Dashboard were the same. However, the Fitbit Dashboard has more severe issues than the ActiVis Dashboard in terms of visibility, recognition, error, and documentation. Still, ActiVis needs improvement compared with Fitbit in terms of control and match, and to solve the catastrophic visibility issue identified.

Figure 13. Usability problems identified in Fitbit and ActiVis dashboards.

RP3: Research Prototype 3

The work on this project is still ongoing. The 3 UX (UX2.1, UX2.2, and UX2.3) of the second ActiVis prototype (RP2) led to new and updated requirements for RP3. Since the last study, the work on this interface has been organized in 2 different branches. The research effort specific to the visualization interface has been split between the different types of users (nurses described in use case 1; family doctors in use cases 2, 3, and 5; and clinician researchers in use cases 4 and 6) with specific charts and interactions but with a common core of data processing functions. The developed visualization prototypes are planned to be integrated into a platform able to read data from different wearable devices available on the market, and integrated into a clinic environment. User evaluations will continue as part of the user-centered design and PD cycles.

Discussion

Principal Findings

The key finding from these PD studies is the derivation from post hoc analysis of nurses' workshop, and the validation by 2 physicians, 1 clinician researcher, and 1 clinician statisticians of the 6 use cases to analyze wearable data for health care professionals. These use cases are assigned to specific user roles: nurses, family doctors, and clinician researchers. They will facilitate the design and development of new data analytics and visualization interfaces to support the particular needs of these users.

UX1

During the PD workshop with nurses evoking their work and relations with patients and other health care professionals, we could not identify specific cultural needs in terms of the visualization of wearable data for health care professionals. Some of the persona and usage scenarios were obviously representative of the local Arabic culture by design, and it is also well-known that particular customs such as prayer times and fasting during the Ramadan Holy month can impact people's patterns of physical activities, sleep, and diet, but none of these aspects finally influenced the more technical use cases we derived from these discussions. The use cases we propose ended up being culturally agnostic (Table 3).

UX2.3

The final evaluation comparing Fitbit and ActiVis dashboards showed there is ample room for improvement even in existing interfaces such as Fitbit, widely available for the general public. We only evaluated use case 1 specific to nurses and already identified some major and catastrophic problems, with severe ratings being more frequent with Fitbit than with ActiVis. Although Fitbit was not necessarily designed to support this use case, it shows that we cannot simply reuse available interfaces to support end users in the best way. Supporting statistical and visual analyses of wearable data from cohorts of patients as stated in use cases targeted at clinician researchers are not optimal or even possible with existing visualization tools and will deserve further investigations.

In general, this project also showed how conducting PD is necessary but still challenging. It has been difficult to plan several of the studies in advance. The use of the opportunistic approach allowed us to use the available local health care professionals throughout the design, development, and validation of RPs presented in this paper. Qatar is a country where 90% of the population are expatriates mixing Western, Asian, and Muslim cultures. Because of the heterogeneous culture and origin of the population, it is challenging to study the levels of health awareness in Qatar [71]. Nevertheless, this is crucial to understand to develop efficient health-targeted visualizations. The population diversity also allowed us to get feedback from non-Qatari, non-Muslim users too. Opening to a wide range of cultures in the same place is of interest to understand what is common or specific to these end users. Although the interface for health care professionals is not impacted by local culture, we know from a previous study [72] that the interfaces involving the patients themselves will need specific care of their local particular health conditions (eg, diabetes or obesity) and Muslim culture (Ramadan Holy month effect on diet, sleep, and physical activity).

Limitations

The study has several limitations. First, a specific set of methods from the user-centered design and PD methodologies was used. Second, the studies were conducted with a selected list of institutions and their experts as participants. Third, several participants were used in each study that was mainly dependent on multiple factors, including availability based on their routine clinical appointments, meetings, and teaching. Fourth, Nielsen's

heuristics were used to diagnose user problems in the prototype that need to be fixed. All these constraints could affect the generalizability of the results. For future studies, we seek higher diversity and a higher number of participants, and domain-specific heuristics to get more generalizable findings.

Conclusion

This paper shows how the use of PD and user-centered design allowed the development of a visualization interface supporting the real needs of health care professionals in Qatar. Although Qatar is an oil-based economy that nurtures a rich multicultural

environment, the use cases we derived from the PD studies happen to be culturally agnostic. We hope these use cases will serve to design future visualization and analytic systems optimized to support the needs specific to nurses, family doctors, and clinician researchers, beyond existing dashboards designed primarily for the general public. This work is still ongoing. A cluster project has now started that is funded under the Qatar National Research Fund [73] and will support further development and integration of these visualizations in a clinical setting to help clinician researchers, doctors, and nurses improve the health of Qatari citizens and residents.

Acknowledgments

We thank Dr Luis Fernandez-Luque for his strong support and thoughtful guidance throughout this research.

Conflicts of Interest

None declared.

Multimedia Appendix 1

Expert comments.

[DOCX File, 23 KB - [humanfactors_v9i2e25880_app1.docx](#)]

Multimedia Appendix 2

Expert recommendations.

[DOCX File, 24 KB - [humanfactors_v9i2e25880_app2.docx](#)]

References

1. Kirwan M, Duncan MJ, Vandelanotte C, Mummery WK. Using smartphone technology to monitor physical activity in the 10,000 Steps program: a matched case-control trial. *J Med Internet Res* 2012 Apr 20;14(2):e55 [FREE Full text] [doi: [10.2196/jmir.1950](#)] [Medline: [22522112](#)]
2. Physical activity. World Health Organization. 2018. URL: <https://www.who.int/news-room/facts-in-pictures/detail/physical-activity> [accessed 2022-11-18]
3. Arora T, Gad H, Omar OM, Choudhury S, Chagoury O, Sheikh J, et al. The associations among objectively estimated sleep and obesity indicators in elementary schoolchildren. *Sleep Med* 2018 Jul;47:25-31. [doi: [10.1016/j.sleep.2018.03.014](#)] [Medline: [29880144](#)]
4. Qatar Biobank. Qatar Biobank Report 2016/2017. Qatar Biobank. 2017. URL: <https://www.qatarbiobank.org.qa/themes/biotech/docs/ANNUAL-REPORT-201617.pdf> [accessed 2021-09-28]
5. Calfas KJ, Long BJ, Sallis JF, Wooten WJ, Pratt M, Patrick K. A controlled trial of physician counseling to promote the adoption of physical activity. *Prev Med* 1996 May;25(3):225-233. [doi: [10.1006/pmed.1996.0050](#)] [Medline: [8780999](#)]
6. King AC, Sallis JF, Dunn AL, Simons-Morton DG, Albright CA, Cohen S, et al. Overview of the Activity Counseling Trial (ACT) intervention for promoting physical activity in primary health care settings. *Activity Counseling Trial Research Group. Med Sci Sports Exerc* 1998 Jul;30(7):1086-1096. [doi: [10.1097/00005768-199807000-00011](#)] [Medline: [9662678](#)]
7. No authors listed. Guidelines for school and community programs to promote lifelong physical activity among young people. National Center for Chronic Disease Prevention and Health Promotion, Centers for Disease Control and prevention. *J Sch Health* 1997 Aug;67(6):202-219. [doi: [10.1111/j.1746-1561.1997.tb06307.x](#)] [Medline: [9285866](#)]
8. Bellg AJ, Borrelli B, Resnick B, Hecht J, Minicucci DS, Ory M, Treatment Fidelity Workgroup of the NIH Behavior Change Consortium. Enhancing treatment fidelity in health behavior change studies: best practices and recommendations from the NIH Behavior Change Consortium. *Health Psychol* 2004 Oct;23(5):443-451. [doi: [10.1037/0278-6133.23.5.443](#)] [Medline: [15367063](#)]
9. Rönkkö K. An Activity Tracker and Its Accompanying App as a Motivator for Increased Exercise and Better Sleeping Habits for Youths in Need of Social Care: Field Study. *JMIR Mhealth Uhealth* 2018 Dec 21;6(12):e193 [FREE Full text] [doi: [10.2196/mhealth.9286](#)] [Medline: [30578186](#)]
10. Sathyanarayana A, Joty S, Fernandez-Luque L, Ofli F, Srivastava J, Elmagarmid A, et al. Sleep Quality Prediction From Wearable Data Using Deep Learning. *JMIR Mhealth Uhealth* 2016 Dec 04;4(4):e125 [FREE Full text] [doi: [10.2196/mhealth.6562](#)] [Medline: [27815231](#)]

11. Cummiskey M. There's an App for That Smartphone Use in Health and Physical Education. *Journal of Physical Education, Recreation & Dance* 2011 Oct;82(8):24-30. [doi: [10.1080/07303084.2011.10598672](https://doi.org/10.1080/07303084.2011.10598672)]
12. Perez-Pozuelo I, Zhai B, Palotti J, Mall R, Aupetit M, Garcia-Gomez JM, et al. The future of sleep health: a data-driven revolution in sleep science and medicine. *NPJ Digit Med* 2020 Mar 23;3(1):42 [FREE Full text] [doi: [10.1038/s41746-020-0244-4](https://doi.org/10.1038/s41746-020-0244-4)] [Medline: [32219183](https://pubmed.ncbi.nlm.nih.gov/32219183/)]
13. Robberechts P. SleepKit: a mobile sleep tracking application for promoting healthy sleep behavior. In: PhD Thesis. Leuven, Belgium: KU Leuven; 2016.
14. Ryokai K. Communicating and interpreting wearable sensor data with health coaches. 2015 May 20 Presented at: 9th International Conference on Pervasive Computing Technologies for Healthcare (PervasiveHealth); May 20-23, 2015; Istanbul, Turkey. [doi: [10.4108/icst.pervasivehealth.2015.259055](https://doi.org/10.4108/icst.pervasivehealth.2015.259055)]
15. Cappers BC, van Wijk JJ. Exploring Multivariate Event Sequences Using Rules, Aggregations, and Selections. *IEEE Trans. Visual. Comput. Graphics* 2018 Jan;24(1):532-541. [doi: [10.1109/tvcg.2017.2745278](https://doi.org/10.1109/tvcg.2017.2745278)]
16. Blascheck T, Besancon L, Bezerianos A, Lee B, Isenberg P. Glanceable Visualization: Studies of Data Comparison Performance on Smartwatches. *IEEE Trans. Visual. Comput. Graphics* 2019 Jan;25(1):630-640. [doi: [10.1109/tvcg.2018.2865142](https://doi.org/10.1109/tvcg.2018.2865142)]
17. Brehmer M, Lee B, Isenberg P, Choe EK. Visualizing Ranges over Time on Mobile Phones: A Task-Based Crowdsourced Evaluation. *IEEE Trans. Visual. Comput. Graphics* 2019 Jan;25(1):619-629. [doi: [10.1109/tvcg.2018.2865234](https://doi.org/10.1109/tvcg.2018.2865234)]
18. Duncan M, Murawski B, Short CE, Rebar AL, Schoeppe S, Alley S, et al. Activity Trackers Implement Different Behavior Change Techniques for Activity, Sleep, and Sedentary Behaviors. *Interact J Med Res* 2017 Aug 14;6(2):e13 [FREE Full text] [doi: [10.2196/ijmr.6685](https://doi.org/10.2196/ijmr.6685)] [Medline: [28807889](https://pubmed.ncbi.nlm.nih.gov/28807889/)]
19. Asimakopoulos S, Asimakopoulos G, Spillers F. Motivation and User Engagement in Fitness Tracking: Heuristics for Mobile Healthcare Wearables. *Informatics* 2017 Jan 22;4(1):5. [doi: [10.3390/informatics4010005](https://doi.org/10.3390/informatics4010005)]
20. Mataraso S, Socrates V, Lekschas F, Gehlenborg N. Halyos: A patient-facing visual EHR interface for longitudinal risk awareness. *bioRxiv* 2019;1:1 [FREE Full text] [doi: [10.1101/597583](https://doi.org/10.1101/597583)]
21. Alagöz F, Ziefle M, Wilkowska W, Valdez AC. Openness to Accept Medical Technology - A Cultural View. Berlin, Germany: Springer; 2011.
22. Mahlich J, Dilokthornsakul P, Sruamsiri R, Chaikyapapruk N. Cultural beliefs, utility values, and health technology assessment. *Cost Eff Resour Alloc* 2018 Jun 1;16(1):19 [FREE Full text] [doi: [10.1186/s12962-018-0103-1](https://doi.org/10.1186/s12962-018-0103-1)] [Medline: [29881328](https://pubmed.ncbi.nlm.nih.gov/29881328/)]
23. Ziefle M, Jakobs EM. New challenges in human computer interaction: strategic directions and interdisciplinary trends. 2010 Jan Presented at: International Conference on Competitive Manufacturing Technologies; January 2010; South Africa p. 389-398 URL: <https://tinyurl.com/bddt2du5>
24. Al-Hazzaa HM, Abahussain NA, Al-Sobayel HI, Qahwaji DM, Musaiger AO. Lifestyle factors associated with overweight and obesity among Saudi adolescents. *BMC Public Health* 2012 May 16;12(1):354 [FREE Full text] [doi: [10.1186/1471-2458-12-354](https://doi.org/10.1186/1471-2458-12-354)] [Medline: [22591544](https://pubmed.ncbi.nlm.nih.gov/22591544/)]
25. Alodhayani AA, Alshaikh OM, Ghomraoui FA, AlShaibani TJ, Algadheeb AS, Bendahmash AW, et al. Correlation between obesity and sleep disturbance in Saudi medical students. *J Phys Ther Sci* 2017 Mar;29(2):181-186 [FREE Full text] [doi: [10.1589/jpts.29.181](https://doi.org/10.1589/jpts.29.181)] [Medline: [28265135](https://pubmed.ncbi.nlm.nih.gov/28265135/)]
26. Alshobaili F, AlYousefi N. The effect of smartphone usage at bedtime on sleep quality among Saudi non- medical staff at King Saud University Medical City. *J Family Med Prim Care* 2019 Jul;8(6):1953-1957 [FREE Full text] [doi: [10.4103/jfmpc.jfmpc_269_19](https://doi.org/10.4103/jfmpc.jfmpc_269_19)] [Medline: [31334161](https://pubmed.ncbi.nlm.nih.gov/31334161/)]
27. Siddiqui AF, Al-Musa H, Al-Amri H, Al-Qahtani A, Al-Shahrani M, Al-Qahtani M. Sleep Patterns and Predictors of Poor Sleep Quality among Medical Students in King Khalid University, Saudi Arabia. *Malays J Med Sci* 2016 Dec 07;23(6):94-102 [FREE Full text] [doi: [10.21315/mjms2016.23.6.10](https://doi.org/10.21315/mjms2016.23.6.10)] [Medline: [28090183](https://pubmed.ncbi.nlm.nih.gov/28090183/)]
28. Vats MG, Mahboub BH, Al Hariri H, Al Zaabi A, Vats D. Obesity and Sleep-Related Breathing Disorders in Middle East and UAE. *Can Respir J* 2016;2016:9673054-9673055 [FREE Full text] [doi: [10.1155/2016/9673054](https://doi.org/10.1155/2016/9673054)] [Medline: [28070158](https://pubmed.ncbi.nlm.nih.gov/28070158/)]
29. Ng SW, Zaghoul S, Ali HI, Harrison G, Popkin BM. The prevalence and trends of overweight, obesity and nutrition-related non-communicable diseases in the Arabian Gulf States. *Obesity Reviews* 2011;12(1):1-13. [doi: [10.1111/j.1467-789x.2010.00750.x](https://doi.org/10.1111/j.1467-789x.2010.00750.x)]
30. Alabdulqader E, Abokhodair N, Lazem S. Human-computer interaction across the Arab world. In: Proceedings of the 2017 CHI Conference Extended Abstracts on Human Factors in Computing Systems (CHI EA '17). 2017 Presented at: 2017 CHI Conference Extended Abstracts on Human Factors in Computing Systems (CHI EA '17); May 6-11, 2017; Denver, CO p. 1356-1359. [doi: [10.1145/3027063.3049280](https://doi.org/10.1145/3027063.3049280)]
31. Alabdulqader E, Lazem S, Nassir S, Saleh M, Armouch S, Dray S. With an eye to the future: HCI practice and research in the Arab world. In: Extended Abstracts of the 2019 CHI Conference on Human Factors in Computing Systems. 2019 Presented at: 2019 CHI Conference on Human Factors in Computing Systems; May 4-9, 2019; Glasgow, UK. [doi: [10.1145/3290607.3299006](https://doi.org/10.1145/3290607.3299006)]
32. Salgado L, Pereira R, Gasparini I. In: Kurosu M, editor. Cultural Issues in HCI: Challenges and Opportunities. Cham, Switzerland: Springer International Publishing; 2015:60-70.

33. Alabdulqader E, Lazem S, Khamis M, Dray SM. Exploring Participatory Design Methods to Engage with Arab Communities. In: Extended Abstracts of the 2018 CHI Conference on Human Factors in Computing Systems. 2018 Presented at: 2018 CHI Conference on Human Factors in Computing Systems; April 21-26, 2018; Montréal, QC, Canada. [doi: [10.1145/3170427.3170623](https://doi.org/10.1145/3170427.3170623)]
34. Bjerknæs G, Bratteteig T. User participation and democracy: a discussion of Scandinavian research on systems development. *Scandinavian Journal of Information Systems* 1995;7(1):73-98.
35. Kanstrup AM, Madsen J, Nøhr C, Bygholm A, Bertelsen P. Developments in Participatory Design of Health Information Technology - A Review of PDC Publications from 1990-2016. *Stud Health Technol Inform* 2017;233:1-13. [Medline: [28125408](https://pubmed.ncbi.nlm.nih.gov/28125408/)]
36. Andersen TO, Bansler JP, Kensing F, Moll J. From Prototype to Product: Making Participatory Design of mHealth Commercially Viable. *Stud Health Technol Inform* 2017;233:95-112. [Medline: [28125417](https://pubmed.ncbi.nlm.nih.gov/28125417/)]
37. Bowen S, McSeveny K, Lockley E, Wolstenholme D, Cobb M, Dearden A. How was it for you? Experiences of participatory design in the UK health service. *CoDesign* 2013 Dec;9(4):230-246. [doi: [10.1080/15710882.2013.846384](https://doi.org/10.1080/15710882.2013.846384)]
38. Byrne E, Sahay S. Participatory design for social development: A South African case study on community-based health information systems. *Information Technology for Development* 2010 Mar 03;13(1):71-94. [doi: [10.1002/itdj.20052](https://doi.org/10.1002/itdj.20052)]
39. Clemensen J, Larsen SB, Kyng M, Kirkevold M. Participatory design in health sciences: Using cooperative experimental methods in developing health services and computer technology. *Qual Health Res* 2007 Jan 01;17(1):122-130. [doi: [10.1177/1049732306293664](https://doi.org/10.1177/1049732306293664)] [Medline: [17170250](https://pubmed.ncbi.nlm.nih.gov/17170250/)]
40. Clemensen J, Rothmann MJ, Smith AC, Caffery LJ, Danbjørg DB. Participatory design methods in telemedicine research. *J Telemed Telecare* 2016 Dec 27;23(9):780-785. [doi: [10.1177/1357633x16686747](https://doi.org/10.1177/1357633x16686747)]
41. Sparud-Lundin C, Josefsson U, Berg M, Hellstrom A, Koinberg I, Nolbris MJ, et al. Use of participatory design in the development of person-centred web-based support for persons with long-term illness. *EJPC* 2013 Nov 18;1(2):369. [doi: [10.5750/ejpc.v1i2.672](https://doi.org/10.5750/ejpc.v1i2.672)]
42. Elovaara P, Iigira FT, Mörtberg C. Whose participation? Whose knowledge?: Exploring PD in Tanzania-Zanzibar and Sweden. In: Proceedings of the Ninth Conference on Participatory Design: Expanding Boundaries in Design. 2006 Presented at: Ninth Conference on Participatory Design: Expanding Boundaries in Design; August 1-5, 2006; Trento, Italy p. 105-114. [doi: [10.1145/1147261.1147277](https://doi.org/10.1145/1147261.1147277)]
43. Chawani MS, Kaasbøll J, Finken S. Stakeholder participation in the development of an electronic medical record system in Malawi. In: Proceedings of the 13th Participatory Design Conference. 2014 Presented at: 13th Participatory Design Conference; October 6-10, 2014; New York, NY p. 71-80. [doi: [10.1145/2661435.2661444](https://doi.org/10.1145/2661435.2661444)]
44. Shidende NH, Iigira FT, Mörtberg CM. An Ethnographically Informed Participatory Design of Primary Healthcare Information Technology in a Developing Country Setting. *Stud Health Technol Inform* 2017;233:131-147. [Medline: [28125420](https://pubmed.ncbi.nlm.nih.gov/28125420/)]
45. Cheng X, Zhao K, Zha X, Du X, Li Y, Chen S, China Health Big Data (China Biobank) project investigators. Opportunistic Screening Using Low-Dose CT and the Prevalence of Osteoporosis in China: A Nationwide, Multicenter Study. *J Bone Miner Res* 2021 Mar 04;36(3):427-435 [FREE Full text] [doi: [10.1002/jbmr.4187](https://doi.org/10.1002/jbmr.4187)] [Medline: [33145809](https://pubmed.ncbi.nlm.nih.gov/33145809/)]
46. Rokach A, Bohadana A, Kotek O, Shuali C, Azulai H, Babai P, et al. Early Detection of COPD: An Opportunistic Case Finding Study in Smokers and Ex-Smokers Visiting a Medical Centre. *COPD* 2021 Jun; Volume 16:1519-1527. [doi: [10.2147/copd.s307483](https://doi.org/10.2147/copd.s307483)]
47. Sheppard D, Clarke E, Hemming K, Martin J, Lilford R. An opportunistic evaluation of a routine service improvement project to reduce falls in hospital. *BMC Health Serv Res* 2021 Jan 22;21(1):79 [FREE Full text] [doi: [10.1186/s12913-021-06073-4](https://doi.org/10.1186/s12913-021-06073-4)] [Medline: [33482793](https://pubmed.ncbi.nlm.nih.gov/33482793/)]
48. Keyworth C, Epton T, Goldthorpe J, Calam R, Armitage CJ. Delivering Opportunistic Behavior Change Interventions: a Systematic Review of Systematic Reviews. *Prev Sci* 2020 Apr 17;21(3):319-331 [FREE Full text] [doi: [10.1007/s11121-020-01087-6](https://doi.org/10.1007/s11121-020-01087-6)] [Medline: [32067156](https://pubmed.ncbi.nlm.nih.gov/32067156/)]
49. Pillay J, Wingert A, MacGregor T, Gates M, Vandermeer B, Hartling L. Screening for chlamydia and/or gonorrhoea in primary health care: systematic reviews on effectiveness and patient preferences. *Syst Rev* 2021 Apr 19;10(1):118 [FREE Full text] [doi: [10.1186/s13643-021-01658-w](https://doi.org/10.1186/s13643-021-01658-w)] [Medline: [33879251](https://pubmed.ncbi.nlm.nih.gov/33879251/)]
50. Saffer D. Designing for Interaction: Creating Innovative Applications and Devices. Indianapolis, IN: New Riders; 2010.
51. Gould JD, Lewis C. Designing for usability: key principles and what designers think. *Commun. ACM* 1985 Mar;28(3):300-311. [doi: [10.1145/3166.3170](https://doi.org/10.1145/3166.3170)]
52. Mao J, Vredenburg K, Smith PW, Carey T. The state of user-centered design practice. *Commun. ACM* 2005 Mar;48(3):105-109. [doi: [10.1145/1047671.1047677](https://doi.org/10.1145/1047671.1047677)]
53. Luna D, Quispe M, Gonzalez Z, Alemraes A, Risk M, Garcia Aurelio M, et al. User-centered design to develop clinical applications. Literature review. *Stud Health Technol Inform* 2015;216:967. [Medline: [26262269](https://pubmed.ncbi.nlm.nih.gov/26262269/)]
54. Brunner J, Chuang E, Goldzweig C, Cain CL, Sugar C, Yano EM. User-centered design to improve clinical decision support in primary care. *Int J Med Inform* 2017 Aug;104:56-64 [FREE Full text] [doi: [10.1016/j.ijmedinf.2017.05.004](https://doi.org/10.1016/j.ijmedinf.2017.05.004)] [Medline: [28599817](https://pubmed.ncbi.nlm.nih.gov/28599817/)]

55. Carr EC, Babione JN, Marshall D. Translating research into practice through user-centered design: An application for osteoarthritis healthcare planning. *Int J Med Inform* 2017 Aug;104:31-37. [doi: [10.1016/j.ijmedinf.2017.05.007](https://doi.org/10.1016/j.ijmedinf.2017.05.007)] [Medline: [28599814](https://pubmed.ncbi.nlm.nih.gov/28599814/)]
56. Risald R, Suyoto S, Santoso AJ. Mobile Application Design Emergency Medical Call for the Deaf using UCD Method. *Int. J. Interact. Mob. Technol* 2018 Jul 20;12(3):168. [doi: [10.3991/ijim.v12i3.8754](https://doi.org/10.3991/ijim.v12i3.8754)]
57. Nguyen KA, Patel H, Haggstrom DA, Zillich AJ, Imperiale TF, Russ AL. Utilizing a user-centered approach to develop and assess pharmacogenomic clinical decision support for thiopurine methyltransferase. *BMC Med Inform Decis Mak* 2019 Oct 17;19(1):194 [FREE Full text] [doi: [10.1186/s12911-019-0919-4](https://doi.org/10.1186/s12911-019-0919-4)] [Medline: [31623616](https://pubmed.ncbi.nlm.nih.gov/31623616/)]
58. Rödle W, Wimmer S, Zahn J, Prokosch H, Hinkes B, Neubert A, et al. User-Centered Development of an Online Platform for Drug Dosing Recommendations in Pediatrics. *Appl Clin Inform* 2019 Aug 07;10(4):570-579 [FREE Full text] [doi: [10.1055/s-0039-1693714](https://doi.org/10.1055/s-0039-1693714)] [Medline: [31390668](https://pubmed.ncbi.nlm.nih.gov/31390668/)]
59. Kozuch K. Apple Watch vs. Fitbit: Which smartwatch brand should you buy? Tom's Guide. URL: <https://www.tomsguide.com/face-off/apple-watch-vs-fitbit> [accessed 2022-09-28]
60. Jary S. Fitbit vs Apple Watch. Tech Advisor. 2021. URL: <https://www.techadvisor.com/feature/wearable-tech/fitbit-vs-apple-watch-3612954/> [accessed 2022-09-28]
61. Fitbit I. Fitbit Reports Third Quarter Results for the Three Months Ended October 3, 2020. Investor. 2020. URL: <https://tinyurl.com/2p8wewvp> [accessed 2021-09-28]
62. Taylor MJ, Vlaev I, Taylor D, Gately P, Ahmedna M, Kerkadi A, et al. A weight-management camp followed by weekly after-school lifestyle education sessions as an obesity intervention for Qatari children: a prospective cohort study. *The Lancet* 2015 Nov;386:S72. [doi: [10.1016/s0140-6736\(15\)00910-1](https://doi.org/10.1016/s0140-6736(15)00910-1)]
63. Aupetit M, Fernandez-Luque L, Singh M, Srivastava J. Visualization of Wearable Data and Biometrics for Analysis and Recommendations in Childhood Obesity. 2017 Presented at: IEEE 30th International Symposium on Computer-Based Medical Systems (CBMS); June 22-24, 2017; Thessaloniki, Greece. [doi: [10.1109/cbms.2017.120](https://doi.org/10.1109/cbms.2017.120)]
64. Rhorbach B. Kreativ nach Regeln – Methode 635, eine neue Technik zum Lösen von Problemen (Creative by rules - Method 635, a new technique for solving problems). *Absatzwirtschaft* 1969;12:73-75.
65. Wieggers K. Seven Truths About Peer Reviews. *Process Impact*. URL: http://www.processimpact.com/articles/seven_truths.html [accessed 2021-09-28]
66. Jacobson I. *Object-Oriented Software Engineering: A Use Case Driven Approach*. Boston, MA: Addison-Wesley; 1992:552.
67. Rogers Y, Sharp H, Preece J. *Interaction Design: Beyond Human-Computer Interaction* (4th edition). Hoboken, NJ: Wiley; 2015.
68. Sears A. Heuristic Walkthroughs: Finding the Problems Without the Noise. *International Journal of Human-Computer Interaction* 1997 Sep;9(3):213-234. [doi: [10.1207/s15327590ijhc0903_2](https://doi.org/10.1207/s15327590ijhc0903_2)]
69. Nielsen J. Ten usability heuristics. Nielsen Norman Group. 1995. URL: <http://www.nngroup.com/articles/ten-usability-heuristics/> [accessed 2022-02-25]
70. Gallello C. UX Check. URL: <https://www.uxcheck.co/> [accessed 2022-11-20]
71. Araújo M, Mejova Y, Aupetit M, Weber I. Visualizing Health Awareness in the Middle East. 2017 Presented at: Eleventh International AAI Conference on Web and Social Media; May 15-18, 2017; Montréal, QC, Canada.
72. Fernandez-Luque L, Singh M, Ofli F, Mejova YA, Weber I, Aupetit M, et al. Implementing 360° Quantified Self for childhood obesity: feasibility study and experiences from a weight loss camp in Qatar. *BMC Med Inform Decis Mak* 2017 May 13;17(1):37 [FREE Full text] [doi: [10.1186/s12911-017-0432-6](https://doi.org/10.1186/s12911-017-0432-6)] [Medline: [28403865](https://pubmed.ncbi.nlm.nih.gov/28403865/)]
73. Abou-Samra AB. Qatar Diabetes Prevention Program. Doha, Qatar: Qatar National Research Fund; 2019. URL: <https://mis.qgrants.org/Public/AwardDetails.aspx?ParamPid=fhgkhjbpc> [accessed 2021-09-28]

Abbreviations

- HMC-Q:** Hamad Medical Corporation-Qatar
- PD:** participatory design
- QCRI:** Qatar Computing Research Institute
- RP:** research prototype
- UX:** user study
- WCM-Q:** Weill Cornell Medicine - Qatar
- WHO:** World Health Organization

Edited by A Kushniruk; submitted 22.11.20; peer-reviewed by L Alsughayer, PCI Pang; comments to author 17.02.21; revised version received 12.12.21; accepted 03.01.22; published 08.04.22.

Please cite as:

Khowaja K, Syed WW, Singh M, Taheri S, Chagoury O, Al-Thani D, Aupetit M

A Participatory Design Approach to Develop Visualization of Wearable Actigraphy Data for Health Care Professionals: Case Study in Qatar

JMIR Hum Factors 2022;9(2):e25880

URL: <https://humanfactors.jmir.org/2022/2/e25880>

doi: [10.2196/25880](https://doi.org/10.2196/25880)

PMID: [35394442](https://pubmed.ncbi.nlm.nih.gov/35394442/)

©Kamran Khowaja, Wafa Waheeda Syed, Meghna Singh, Shahrad Taheri, Odette Chagoury, Dena Al-Thani, Michaël Aupetit. Originally published in JMIR Human Factors (<https://humanfactors.jmir.org>), 08.04.2022. This is an open-access article distributed under the terms of the Creative Commons Attribution License (<https://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Human Factors, is properly cited. The complete bibliographic information, a link to the original publication on <https://humanfactors.jmir.org>, as well as this copyright and license information must be included.

Original Paper

A Nurse-Led Multimedia Intervention to Increase Patient Participation in Recovery After Knee Arthroplasty: Hybrid Type II Implementation Study

Jo McDonall¹, PhD; Bernice Redley^{1,2}, PhD; Patricia Livingston¹, PhD; Ana Hutchinson^{1,3}, PhD; Richard de Steiger^{4,5}, MBBS, PhD; Mari Botti^{1,3}, PhD

¹School of Nursing and Midwifery, Faculty of Health, Deakin University, Geelong, Australia

²Centre for Quality and Patient Safety Research-Monash Health Partnership, Deakin University, Geelong, Australia

³Centre for Quality and Patient Safety Research-Epworth HealthCare Partnership, Geelong, Australia

⁴Epworth HealthCare, Epworth Victor Smorgon Chair of Surgery, Department of Surgery, Richmond, Australia

⁵The University of Melbourne, Melbourne, Australia

Corresponding Author:

Jo McDonall, PhD

School of Nursing and Midwifery

Faculty of Health

Deakin University

1 Gheringhap St.

Geelong, 3220

Australia

Phone: 61 92446630

Email: jo.mcdonall@deakin.edu.au

Abstract

Background: Advances in digital technology and the use of multimedia platforms to deliver information provide clinicians with a unique opportunity to develop innovative ways to consistently provide high-quality, accessible, and evidence-based information to support patient participation. Introducing new technologies into everyday acute care clinical practice can be difficult.

Objective: The aim of this paper was to provide a description of an implementation strategy and the subsequent evaluation undertaken to examine the contextual factors important to the successful adoption of new technology by nurses in the context of acute postoperative care.

Methods: Implementation of the intervention and process evaluation was undertaken in 3 phases: phase 1, preimplementation stakeholder engagement and identification of barriers and enablers to implementation; phase 2, supported implementation of the intervention; and phase 3, evaluation of uptake, usability, and acceptability of the intervention in clinical practice.

Results: The outcomes of the implementation of the multimedia intervention in the context of acute postoperative care were positive. Of the 104 patients in the intervention group, 103 (99%) received the intervention. All 103 patients completed the 8-item intervention questionnaire and 93.3% (97/103) were interviewed on day 3 to evaluate usability, uptake, and acceptability. Of these 97 patients, almost all (n=94, 91%) found the program easy to use and most (n=64, 62%) could view the MyStay Total Knee Replacement program as often as they wanted. The findings also suggest that the time to implement the program was minimal (5-10 minutes). Collaboration with nurses and patients before and during implementation to identify potential barriers to successful implementation of the intervention was essential to develop timely strategies to overcome these barriers. To ensure end-user engagement, careful consideration was given to nurses' views on who was responsible for facilitating this intervention.

Conclusions: The findings provide evidence that the structured implementation of the multimedia intervention was robust and successful in terms of patient participant recruitment and application; however, it was difficult to assess the level of engagement by nurse clinicians with the program.

Trial Registration: Australian New Zealand Clinical Trials Registry (ANZCTR) ACTRN12614000340639; <https://anzctr.org.au/Trial/Registration/TrialReview.aspx?ACTRN=12614000340639>

KEYWORDS

patient participation; multimedia; nurse-facilitated; knee arthroplasty; orthopedic surgery; acute care; nurse; participatory medicine; digital technology

Introduction

Background

Advances in digital technology and the use of multimedia platforms to deliver information provide clinicians with a unique opportunity to develop innovative ways to consistently provide high-quality, accessible, and evidence-based information to support patient participation. To date, multimedia tools have been used in a wide range of health situations, including providing preoperative education to prepare patients to undergo specific procedures and provide consent [1-7]; assisting patients to make informed decisions regarding treatment [8,9]; enabling self-management in chronic illness [10]; supporting postoperative care (eg, how to use a patient-controlled analgesic pump after surgery) [11]; and improving patient overall satisfaction [12]. There is emerging use of technology to facilitate patient participation in acute postoperative contexts. The effectiveness of using digital, multimedia platforms to enhance patient participation in their care is directly affected by nurses' attitudes [13]. If nurses perceive that a new technology will be burdensome, unreliable, or does not fit into their workflow, they are less likely to promote its use by their patients. Implementation strategies that specifically target the range of individuals involved in delivering patient care and organizational processes are needed to successfully introduce and embed novel technologies and interventions into clinical practice [14,15].

A novel multimedia intervention, *MyStay Total Knee Replacement (MyStay TKR)*, was developed specifically for use by patients after undergoing total knee replacement by Enlighten Health, a medical multimedia production company specializing in validated content for patient and clinical education. *MyStayTKR* was developed using an iterative, multi-method approach aimed at ensuring that program content was valid and reflected an optimal balance among procedure-specific best evidence, current clinical practice, and patient preferences. *MyStay TKR* was designed to be both nurse-facilitated and patient self-directed; that is, accessed and used independently by patients as a stand-alone program packaged for iPad (Apple Inc) presentation [16]. The intervention has two interacting components: (1) information tailored to each day of recovery to enhance patients' understanding of their goals of recovery and (2) explicit information outlining how to achieve their recovery goals. Multimedia through iPad technology was selected as the intervention most likely to be effective in influencing patient participation in the context of acute postoperative recovery because it places minimal burden on nurses and patients, has continuous availability, and is intuitive and easy to use [17]. The multimedia intervention was designed to deliver information that was explicit, actionable, nonambiguous, and tailored specifically to the daily goals that support patient recovery after total knee replacement surgery (Figures 1-3).

Figure 1. MyStay Total Knee Replacement landing page examples.



Figure 2. MyStay day 1 after surgery My Activity page.

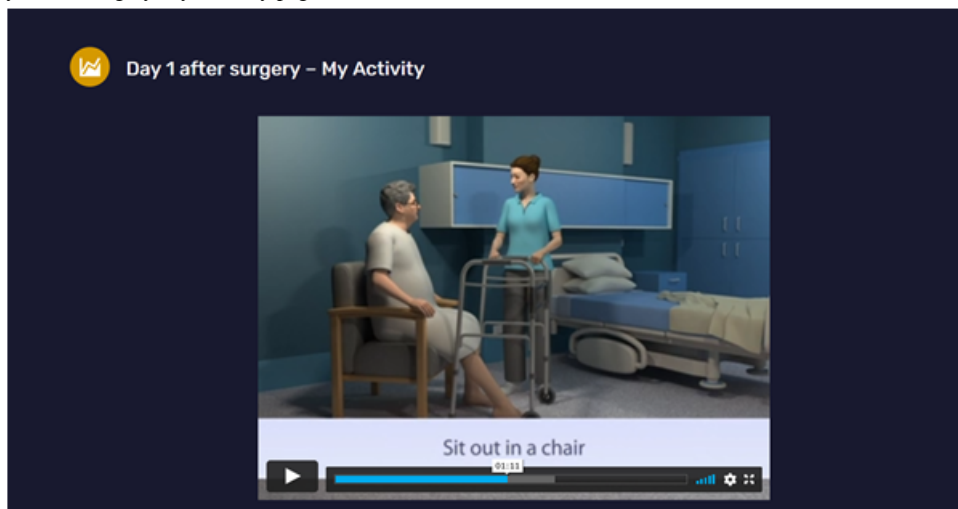
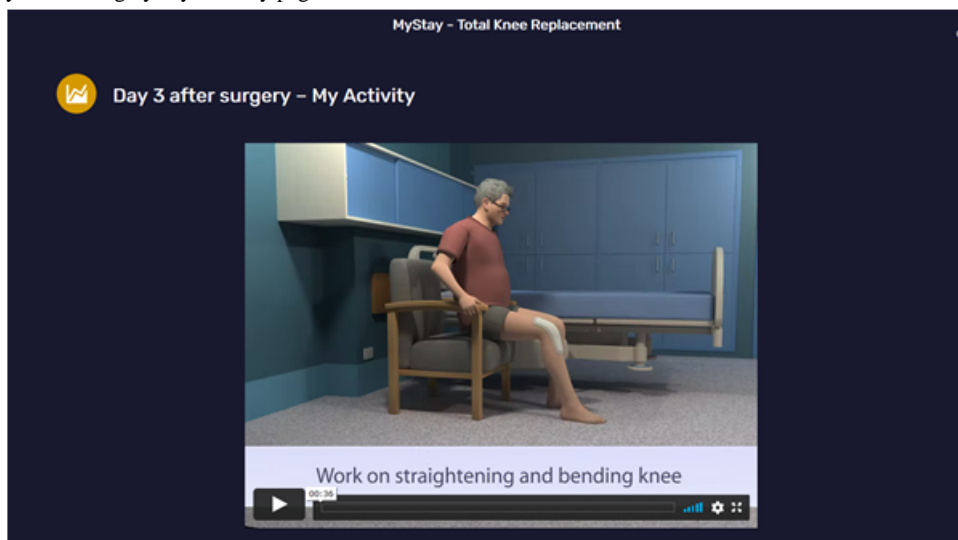


Figure 3. MyStay day 3 after surgery My Activity page.



Objectives

This paper reports an implementation strategy and evaluation undertaken to examine the contextual factors important to the successful adoption of *MyStay TKR* by nurses and patients in the context of acute postoperative care. A cluster randomized controlled trial was conducted to test the effectiveness of the intervention and is reported elsewhere [16]. The implementation strategy and evaluation reported here were designed to assess the degree to which each element of the program was implemented. The specific objectives of this evaluation were to determine the following:

1. The system or environmental factors that may have had an impact on the implementation of the intervention.
2. Patient usability, uptake, and acceptability of the multimedia intervention in the context of acute recovery after surgery.

Methods

Implementation and Process Evaluation

A prospective concurrent implementation evaluation of the *MyStay TKR* intervention was undertaken. The intervention was implemented using a structured standardized approach with boundaries established to limit variation [18]. Consistent evidence-based implementation processes were used on the wards and involved multiple methods [19]. The implementation and process evaluation were undertaken in 3 phases: phase 1, preimplementation stakeholder engagement and identification of barriers and enablers to implementation; phase 2, supported implementation of the *MyStay TKR* intervention; and phase 3, evaluation of patient usability, uptake, and acceptability of the *MyStay TKR* multimedia tool in clinical practice. The implementation strategies used and data collected to measure the effectiveness of the implementation process in each phase of the study are summarized in Table 1.

Table 1. Methods and data collection across 3 stages of the MyStay Total Knee Replacement (MyStay TKR) trial implementation.

Trial phase	Implementation strategies used	Evaluation methods
Phase 1: preimplementation (nurses)	<ul style="list-style-type: none"> • Purposive group interview • Ward meetings and in-service education • Flyers and handouts • Email correspondence 	Analysis of interview and meeting data
Phase 2: implementation (nurses and patients)	<ul style="list-style-type: none"> • Daily ward visits (intervention and control wards) • Daily field observations • One-to-one and ward meetings • Handouts and flyers • Correspondence through patients' bedside whiteboards 	Analysis of meeting notes and observation field notes using qualitative content analysis
Phase 3: evaluation		
Nurses		
	<ul style="list-style-type: none"> • Observations of clinical practice and incidental staff feedback 	Analysis of field notes
Patients		
	<ul style="list-style-type: none"> • Uptake and use of the MyStay TKR intervention by patients • Patient self-reported questionnaires 	Analysis of interview data and descriptive statistics using SPSS software (version 23; IBM Corp)

Phase 1: Stakeholder Engagement

Leadership Engagement

Before implementation of the *MyStay TKR* intervention, nurses (n=4) were purposively sampled to participate in a group interview. All 4 participants were permanent staff employed on the orthopedic ward and included 1 (25%) educator, 2 (50%) senior registered nurses, and 1 (25%) graduate nurse. The focus of the discussion was how best to embed the multimedia intervention into nurses' everyday practice and to identify strategies to mitigate nurses' perceived barriers to successful implementation.

Staff Education Sessions

A range of activities were undertaken before the commencement of the trial to support successful implementation of the *MyStay TKR* intervention; these included informing all surgeons, nurses, and physiotherapists about the study and expectations of their involvement, as well as ensuring that the clinicians were exposed to the intervention and familiar with content and navigating the program on the iPads. At ward and in-service meetings, the project was described in detail along with a demonstration of the animation intervention designed for patients and any questions were addressed. In total, 3 formal ward meetings were conducted before implementing the program on each ward, and additional in-service meetings were held on each ward until >80% of the ward nursing staff had received training in using the *MyStay TKR* tools. In addition to the daily meetings, 1 *night* meeting was held on each ward to ensure that the permanent

night staff were also well informed about the study. Orthopedic surgeons and physiotherapists were also involved in information sessions related to the study to ensure their support for the study.

Development of Promotional Materials

Handouts and flyers were developed and placed at the nurses' stations and in break rooms to engage nurses and inform them of the study. On each flyer the researchers' contact details were provided to invite questions or suggestions. Nurse unit managers, physiotherapists, and ward nursing staff were sent regular emails to provide updates on the stages of the study throughout the trial period.

Phase 1: Data Collection

The semistructured interview with key clinical nurse leaders was recorded and transcribed for later analysis to identify barriers and facilitators to implementation of the *MyStay TKR* intervention. During the staff education sessions, field notes captured reports of barriers and the suggested strategies identified during each of these encounters.

Phase 2: Supported Implementation of the MyStay TKR Intervention

Overview

Application of the intervention procedure involved a structured process that included promotion and awareness raising, patient engagement, and development of tailored solutions to implementation barriers. These processes are outlined in [Table 2](#).

Table 2. Application of the intervention procedure.

Key process	Procedure
Identification of patients enrolled in the trial	<ul style="list-style-type: none"> At the beginning of each day shift, nurse unit managers, and associate nurse unit managers were informed of the following: <ul style="list-style-type: none"> The researcher presence on the ward A list of patients enrolled in the trial on their ward identifying the “day” after surgery The exact number of iPads required per ward per day and ensure that they were charged and ready for use
Application of intervention procedure	<ul style="list-style-type: none"> Identify the nurse responsible for the care of patient participants Confirm with the nurse that the patient is enrolled in the study and will need to view the iPad animation Identify day 1 patients and provide and secure the iPad and explain how to use the device and navigate the program Patients instructed to watch the animation on the iPad and call their nurse once they have finished to discuss the content The nurse will confirm and clarify any questions the patients may have regarding the information The iPads remain with the patient for the duration of their stay Laminated flyer attached to the patients’ medical record and note on the individual patients’ communication board were used to remind clinical staff that the patient was in the study Patients’ nurses are responsible for ensuring that the iPad is charged overnight Telephone call to the wards nightly at 10 to remind them to charge the iPads
Strategies used throughout the trial to maintain engagement by nursing staff	<ul style="list-style-type: none"> One-to-one discussions among ward nurses, physiotherapists, surgeons, and the nurse researcher Telephone calls to associate nurse unit managers on afternoon shifts at 8 PM each day to ask that they remind staff to charge the iPads overnight Regular attendance at ward meetings by the nurse researcher where questions could be answered and strategies discussed to assist with the implementation Laminated cards were placed in patient notes, and a sign was placed on the whiteboard above the patient bed area Patients themselves reminded staff to attend to the iPad; for example, to plug in the iPad for charging overnight

Promotion and Awareness Raising

To support successful implementation of the *MyStay TKR* intervention, the researcher (JM) conducted daily ward visits for the duration of the trial. The intent of these visits was to promote uptake of the *MyStay TKR* intervention and to support adherence to key processes of care. At these ward visits, the researcher ensured that any casual staff were familiar with the trial, placed flyers in patients’ notes and on bedside whiteboards to alert clinical staff that the patient was enrolled in the trial, obtained ongoing feedback from the health care team about any barriers to implementation of the intervention, and observed practices related to implementation and usability of the intervention (patient and clinician engagement).

Patient Engagement and Facilitation of Patient Participation

During the implementation phase, patients together with the researcher navigated the *MyStay TKR* animation on the iPad. Each section of the program was explained until patients were comfortable with access and could follow the program. This introduction to the program took 5 to 10 minutes, depending on the patient’s familiarity with the iPad. The iPad was then left with the patients, who were informed that they could use the program as often as they wished. Patients were also instructed to call their nurse to inform them that they had finished watching the program. The nurse would then clarify any questions the patient may have regarding the information

provided, and it was anticipated that a discussion regarding the goals of the day would ensue. Physiotherapists are crucial in mobilizing patients after surgery and restoring mobility in the knee joint. As such, physiotherapists were involved in the development of *MyStay TKR* content through the Delphi process and reviewing of the exercise components of the multimedia, as well as subsequently during implementation of *MyStay TKR*. During implementation, physiotherapists encouraged patients to view the *MyStay TKR* modules on the iPads. This was achieved during their initial visit on day 1 with patients and also throughout the patients’ stay during follow-up visits. Physiotherapists asked patients to watch the exercise component of *MyStay TKR* in their absence.

Tailored Solutions

During the implementation phase, the researcher obtained ongoing feedback from staff and patients regarding barriers and facilitators to implementation and use of the *MyStay TKR* intervention. To support successful implementation of *MyStay TKR*, the researcher worked with the nursing team based on the ward to identify time-sensitive solutions to these barriers.

Phase 3: Usability, Uptake, and Acceptability of the MyStay TKR Intervention

The effectiveness of the implementation strategy was evaluated using data on the uptake and use of the *MyStay TKR* by patients as well as by obtaining patient feedback about the uptake, usability, and acceptability of using the program to support their

recovery. All patients who were randomized to an intervention ward were given the *MyStay TKR* evaluation questionnaire, an 8-item self-report tool specific to the intervention. This questionnaire was designed to uncover the ease of use, satisfaction with, and effectiveness of the multimedia program to aid in the patients' recovery.

Data Analysis

The Theoretical Domains Framework was used to inform the identification of barriers and facilitators to practice and behavior change with regard to both clinicians and patients [20].

There were three components to the analysis of the study data:

1. Qualitative content analysis was performed to identify key themes that emerged from the focus group discussion and the staff education sessions. The transcripts were independently reviewed for factual content by 2 researchers (JM and MB), who formed agreement on key emergent codes for thematic analysis. The codes were then grouped to identify key themes and subthemes. All members of the research team reviewed the identified themes and subthemes, and the thematic structure was determined by consensus.
2. Throughout the implementation phase the researcher collected field notes describing these communications, and any observations made by the researcher related to implementation of the intervention were transcribed in a field diary. These notes were coded for recurring themes in terms of barriers and facilitators.

3. Descriptive statistics were used to analyze the uptake and use of *MyStay TKR* as well as patient responses to the *MyStay TKR* evaluation questionnaire.

Ethics Approval

The cluster randomized controlled trial and the implementation evaluation were approved by the health service and university institutional ethics committees (Epworth HealthCare Human Research and Ethics Committee, 598-13, and Deakin University Human Research Ethics Committee, 2013-195).

Results

The study results are described in 3 sections to reflect each phase of the implementation evaluation.

Outcomes of Phase 1

Overview

In total, 3 themes were derived from the analysis of the transcripts of the clinician group interview, ward meetings, and one-to-one communications with nurses, and these were used to inform how to embed the intervention into everyday practice on the wards. The themes were as follows: (1) the potential burden of introducing the intervention for staff, (2) perceived difficulties associated with the age of patients and ease of use of technology, and (3) concerns about safety and security of the iPad within the ward (Table 3). There were no concerns raised by physiotherapists or surgeons regarding application of the intervention in the preimplementation stage.

Table 3. Perceived barriers to implementation and strategies used to address these barriers.

Barrier or concerns identified by nurses	Illustrative quote	Strategies used to address concerns or barriers
Potential burden of introducing the intervention for nursing staff	<ul style="list-style-type: none"> “Can you guarantee this [iPad intervention] will not increase our already busy workload? I mean, if we have to spend time going through this iPad [intervention] then it’s going to make it harder for us isn’t it...I mean, we just don’t have the time.” [Nurse ID 2] “I don’t know, I think there’s a lot going on in the morning...we [the nurses] are busy and flat out. First thing it is probably easier if someone else does it [goes through the program with the patient] and not leave it up to the nurses?” [Nurse ID 1] 	<ul style="list-style-type: none"> Implementation of the intervention on day 1 of patients’ recovery was carried out by the researcher to ensure that patients could use the iPad and navigate the program Patients who were classified as postoperative day 1 received an explanation of the iPad and navigation after handover and before breakfast, at approximately 8 AM each day
The age of patients and ease of use of technology	<ul style="list-style-type: none"> “With the older patients we may have to teach them how to use the iPad [intervention] or they may not be able to use it at all. Do you think this is very realistic, I mean for them to use it?” [Nurse ID 3] “Yes, some of them have other comorbidities, you know, such as arthritis, it may be harder for them...we will have to push it for them? If that’s the case, I don’t think we will have the time.” [Nurse ID 4] “I don’t think it should be an issue, my grandparents have one and they use it ok.” [Nurse ID 1] 	<ul style="list-style-type: none"> A flyer to assist patients to navigate the program themselves was provided to all patients Once patients were familiar with the iPad the nurses felt that they were able to focus on the content of the program
Security and safety of the equipment and infection control	<ul style="list-style-type: none"> “So where are you going to put it [iPad intervention]? You don’t want it to get in the way. There’s not much room anyway with all their [patients] stuff. Perhaps it could be put on the bedside tables so we can get it out of the way if we need to?...What about keeping it clean, what do you think?...Have you thought about the cross contamination?” [Nurse ID 1] “Yes, you have to make sure it doesn’t walk either...if it’s not secure, things walk here, how will you make sure it stays with the patient? And what about if it gets dropped they are very sensitive these iPads...what will happen there...do you have lots of replacements?” [Nurse ID 3] 	<ul style="list-style-type: none"> To address security concerns, the iPad was secured to each patient’s movable bedside table with a locked cable Each iPad was secured inside a locked tough case that was drop-, smash-, and splash-proof The infection control nurse approved the cleaning protocol for each iPad before transfer to another patient. Wiping the iPad and all associated material (cords, case, etc) with an alcohol-impregnated cloth was approved as sufficient cleaning between patients Cleaning occurred on collection of the iPad when a participating patient was discharged from hospital

Potential Burden of Introducing the Intervention for Nursing Staff

Nurses expressed concern that they may need to facilitate the use of the iPad and assist patients to navigate the program the first time they were exposed to *MyStay* TKR. They thought that this would take a significant amount of time, particularly during the busy morning period that includes clinical handover and patient assessment. There was also worry that there may be an additional burden on nursing staff during the patients’ stay where they may have to reintroduce and reiterate aspects of the *MyStay* TKR content with patients each day, thus increasing their workload.

The Age of Patients and Ease of Use of Technology

There were mixed attitudes regarding the age of the patients and their ability (physical and mental) to use the iPad. Some

(2/4, 50%) of the nurses indicated that older patients may be unfamiliar with portable devices or unable to use them.

Safety and Storage of Equipment

Nurses were worried about the physical location of the iPad in patients’ rooms and stated that the device could add to existing clutter and be removed or stolen or dropped and broken. The potential for cross-contamination and risk of infection was also raised.

Outcomes of Phase 2

Overview

Table 4 outlines the patient participant characteristics at baseline. During the implementation phase, the strategies outlined in Table 3 were applied to address the potential barriers to successful implementation.

Table 4. Patient participant baseline characteristics (N=104).

Characteristics	Values
Age (years), mean (SD)	65.25 (9.77)
Sex, n (%)	
Male	40 (38.5)
Female	64 (61.5)
Living arrangements, n (%)	
Living communally	88 (84.6)
Living alone	16 (15.4)
Marital status, n (%)	
Partnered	84 (80.8)
Not partnered	10 (9.6)
Widowed	10 (9.6)
Country of birth, n (%)	
Australia	76 (73.1)
United Kingdom	11 (10.6)
Other	8 (7.7)
Europe	6 (5.8)
Asia	2 (1.9)
New Zealand	1 (0.9)
Language spoken at home (primary), n (%)	
English	102 (98)
Italian	1 (0.9)
Other	1 (0.9)
Employment status (preadmission), n (%)	
Retired	52 (50)
Full time	24 (23.1)
Part time or casual	16 (15.4)
Unemployed	7 (6.7)
Other	5 (4.8)

Of the 104 participants recruited for the intervention group, only 1 (0.9%) patient was unable to receive the multimedia intervention in the trial. This deviation was due to factors outside the control of the study: the patient had a serious postoperative complication and therefore was unable to receive the intervention. In total, 94.2% (97/103) of the patient participants were interviewed on day 3. Reasons for participants not interviewed were as follows: too unwell, not available on the ward at the time, or declined to be interviewed. Interview duration ranged from 12 to 75 minutes. Most (94/97, 97%) of the interviews were conducted between 9 AM and 2 PM at the patients' bedside; the rest (3/97, 3%) of the interviews were conducted at a later time (after 5 PM) at the patients' request.

During the interviews, patients reported a range of structural, clinician-related, and patient-related barriers to use of the *MyStay TKR* program. These barriers were addressed as they were identified as described in the following sections.

Structural Factors

The physical location of the iPads presented a problem when trying to ensure that the program was always available for patients when they wanted to access it. Because of physical constraints of space, several options were tested until agreement was reached about the ideal location. Initially, the iPads were secured to the patients' bedside trolleys to enable the iPad to be moved around if patients decided to sit out of bed; however, this caused problems for the food services staff who found it difficult to find room to place patients' food trays. The decision to move the iPads to the patient's bedside locker was made in consultation with the patients as well as food services and nursing staff. The cord that tethered the iPad to the bedside table was long enough for the iPad to be placed on the bed should patients decide to sit out of bed and view the presentations. On several occasions nurses and the services staff moved the patients' iPads to the back wall "to keep it out of the way." This

then prevented patients from watching the iPad as they could not reach it. These prohibitive behaviors were fed back to the nurses caring for the patient on the day.

Clinician-Related Factors

Nurses' attitudes toward the program were critical to its successful implementation. In the third week after commencement of the project, 3% (3/97) of the patients commented that nurses (n=2) had stated that they were "sick of these iPads" and "these iPads just get in the way." These comments can influence patients to question the use of the program and can negatively affect their confidence to ask nurses questions related to the program. To address these issues, group and one-to-one discussions were held with the nursing staff to determine what strategies might be implemented to overcome these perceptions. Field notes revealed that of the 103 iPads used over a period of 14 months, there were 17 (16.5%) with flat batteries; however, the majority (n=13, 76.5%) of these were in the initial rollout phase. Reasons for the flat batteries outlined by nursing staff were "forgot to put on charge," "no charger available," "needed the charging plug for another appliance," and "unable to charge" (2/103, 1.9% of iPads were *missing* the charging adapter). Throughout the trial period this practice improved, with only 3.9% (4/103) of the iPads noted with flat batteries after approximately 1 month following commencement of the trial.

Patient-Related Factors

Difficulties encountered by patients in using the iPad included being unable to watch the entire program because of sleepiness or tiredness, difficulty remembering to watch the program, and being too unwell to watch because of pain or other complications. Strategies were discussed with each patient and their nurse during the daily visit, and methods to overcome

barriers to use were agreed. For example, the patients who were too tired to watch all of the program at once were directed to watch only small clips at a time and nurses would remind them to watch more throughout the day. If patients were in pain, they were reminded by nurses to watch the program later in the day. No barriers were identified by patients in relation to the information delivery using the iPad.

Outcomes of Phase 3: Uptake, Usability, and Acceptability of the MyStay TKR Intervention

Of the 104 recruited participants, only 1 (0.9%) randomized patient was unable to receive the multimedia intervention in the trial because they experienced a serious postoperative complication (cerebrovascular accident) and were therefore unable to receive the intervention.

All 103 patients completed the 8-item intervention questionnaire on day 3. Overall, 66% (68/103) of the patients reported that they had viewed the iPad program more than once in the previous 24 hours, 29.1% (30/103) had viewed the program once, and 4.8% (5/103) reported that they had not viewed the program in the previous 24 hours. Reasons for not viewing the iPad program were as follows: watched the entire program on days 1 and 2 after the surgery, unable to view because of illness, too tired to watch at the time, and they planned to watch the program later in the day.

Almost all (94/103, 91.3%) patients found the program easy to use. In total, 62.1% (64/103) of the patients reported that they felt they could view the program as often as they wanted. Reasons for not viewing the program as often as they would have liked included feeling too tired or too unwell (24/33, 73%), technical issues with the iPad having a flat battery (11/33, 33%), and concerns about the voiceover on the program disturbing patients in shared rooms (1/33, 3%; [Table 5](#)).

Table 5. Patients' reasons for not viewing the program on the iPad as often as they wanted (some patients indicated multiple reasons; N=39, 38%).

Reason stated for not viewing the program as often as wanted	Values, n (%)
Too tired (including visitors)	12 (36)
Too unwell (predominately nausea)	12 (36)
iPad did not work properly when I had the opportunity to watch (battery flat)	11 (33)
No time, (patient) too busy	7 (21)
Pain too severe	6 (18)
iPad not available when I had the opportunity (not in reach)	4 (12)
Forgot about watching it	4 (12)
Did not understand the content	2 (6)

As the intervention was designed to be nurse-facilitated, patients were asked on day 3 whether the nurses responsible for their care had discussed the program with them in the previous 24 hours. Only 21.4% (22/103) of the patients reported that nurses had discussed the program with them in the previous 24 hours.

Patients' reported satisfaction with the intervention was high, as reflected in a mean score of 8.63 (SD 2.05) out of 10. No problems with navigation of the program on the iPad were reported.

Discussion

Principal Findings

This evaluation of whether a multimedia intervention delivered through an iPad could be successfully implemented on acute orthopedic wards established that most (94/103, 91.3%) of the patients found the program easy to use, with their reported satisfaction with the intervention being high (mean score of 8.63, SD 2.05, out of 10), and the program required minimal time for orientation. Collaboration with clinicians and patients

before and during implementation to identify potential barriers to successful implementation of the intervention was essential to develop timely strategies to overcome these barriers. To ensure end-user engagement, careful consideration was given to nurses' views on who was responsible for facilitating this intervention. Several methods were adopted to ensure that nurses had the opportunity to discuss concerns and express their opinions about embedding the implementation into their everyday clinical practice. The effects on nursing staff workload, the physical location of the iPad, and the safety and security of the device were identified as key areas of concern and were addressed in the implementation plan.

The intervention was implemented using a structured, standardized, and evidence-based approach [18,19]. The intervention was designed to be delivered in the context of usual care delivery; however, nurses were reluctant to perform the initial orientation with patients because of concerns that instructing patients on the use and navigation of the iPads would be time consuming and would interfere with the provision of patient care.

On day 3 after the surgery, patients reported low levels of nurse engagement with the intervention. There are several possible explanations for this: it is possible that nurses were satisfied that patients were engaging sufficiently with the intervention or that there had been higher levels of interaction during the previous postoperative days. It is possible that nurses were not engaging with the *MyStay TKR* program and did not see it as a tool to set goals of care with patients to assist them with their recovery. The challenge for future studies is to demonstrate to nurses that these types of interventions will not have an impact on their workloads [18]. In fact, the time needed to explain the program to patients was very brief and could easily be incorporated into everyday clinical practice.

Embedding interventions into clinical practice has been reported to be challenging, particularly in the acute care setting where work is often fast paced, with nurses caring for patients who are acutely ill after surgery [20]. Implementation of the intervention in this study required nurses to facilitate interactions among themselves, the multimedia program, and patients to create opportunities for patients to discuss their goals of recovery and negotiate pain management. This element required a patient-centered approach [21-23], which can be difficult to achieve in practice when nurses perceive that their workload is already high. Several acute care studies have reported that nurses spend only a small amount of time with each patient [24-27], and the acuity levels of the patients in the postoperative context also result in some patients being allocated more time than others [28].

As with any new technology designed for patients in the clinical setting, ease of use is a primary design consideration. Most (64/103, 62.1%) of the patients reported that they were able to view the program as often as they liked without restriction. Patients also successfully navigated the program independently, and all (103/103, 100%) patients interacted with the program at least once a day. However, the patients' acuity levels did limit their level of interaction. These findings are consistent with those of other studies that have evaluated the

implementation of a multimedia intervention in acute care [17,29-31].

Reasons stated by patients for not interacting with the *MyStay* program were predominately related to the acuity of their illness rather than the program itself, suggesting that usability was not a problem. Consistent with the findings of Cook et al [29], the major barrier for patients in engaging with the *MyStay* program was tiredness and nausea, both common symptoms in the acute postoperative period. An advantage of the program being available 24 hours a day was that patients could access the program when it suited them. In previous studies where patients had limited access to interventions, usability was compromised [12]. A study by Chu et al [32] reported that 71% of patient time in hospital was considered *downtime*; that is, patients were not occupied with diagnostic tests or other activities. This suggests that there is ample opportunity for patients to engage with an intervention program throughout the day if there is flexibility in availability. An additional advantage of the multimedia platform is that patients' families could also view these programs during their visits to help to reinforce the goals of recovery.

Nurses' concerns that older age may hinder patients' ability to use the iPad technology was not identified as a limiting factor in this study. Of the 103 patients, only 2 (2%) stated that they were *computer illiterate* and that this was a reason why the program was not easy to use. Advanced age was not identified as a factor affecting usability; indeed, a patient aged 95 years found the iPad so usable that he indicated he would purchase one when he was discharged. Our findings are similar to those of Cook et al [29] who found that patients can in fact interact with a multimedia device, regardless of age: 91.3% (94/103) of the patients reported it to be easy to use; reasons for the patients (9/103, 8.7%) who indicated difficulty included flat battery, lack of concentration because of health, or the sound was poor. The majority of these factors were rectified during the trial.

Creating an opportunity for patient participation without placing an additional burden on clinicians and patients was considered critical in this study. The *MyStay TKR* intervention was designed to be easily navigated by patients and nurses in the acute care environment [33]. Time spent by the researcher orientating patients to the technology was 5 to 10 minutes initially and then 2 to 5 minutes per day with individual patients. It is concluded therefore that the *MyStay TKR* intervention can be incorporated into everyday routine care, despite the acuity of the environment and the time required for nurses to allocate in applying (not facilitating) the program is low and feasible [32]. These findings are consistent with those of other studies that have implemented multimedia interventions for patients in hospital [17,34,35].

Conclusions

Implementation of a nurse-led multimedia intervention to increase patient participation in recovery after total knee replacement was achievable. The findings demonstrated that the implementation of the *MyStay TKR* multimedia intervention was robust and structured and successful in terms of patient participant recruitment and application; however, it was difficult to assess the level of engagement by nurse clinicians with the program. Furthermore, the findings indicate that a multimedia

program designed as a platform to promote patient participation within acute care environments that can present challenges to engagement is feasible and is associated with high patient satisfaction.

Acknowledgments

This research was supported by funding from philanthropic organizations.

Data Availability

Evaluation data are available upon request.

Conflicts of Interest

None declared.

References

1. Armstrong AW, Alikhan A, Cheng LS, Schupp C, Kurlinkus C, Eisen DB. Portable video media for presenting informed consent and wound care instructions for skin biopsies: a randomized controlled trial. *Br J Dermatol* 2010 Nov;163(5):1014-1019. [doi: [10.1111/j.1365-2133.2010.10067.x](https://doi.org/10.1111/j.1365-2133.2010.10067.x)] [Medline: [20977443](https://pubmed.ncbi.nlm.nih.gov/20977443/)]
2. Sowan AK, Beraya AR, Carrola A, Reed C. Effect of a multimedia patient decision aid to supplement the informed consent process of a peripherally inserted central venous catheter procedure: pre-post quasi-experimental study. *JMIR Med Inform* 2018 Nov 28;6(4):e11056 [FREE Full text] [doi: [10.2196/11056](https://doi.org/10.2196/11056)] [Medline: [30487117](https://pubmed.ncbi.nlm.nih.gov/30487117/)]
3. Beamond BM, Beischer AD, Brodsky JW, Leslie H. Improvement in surgical consent with a preoperative multimedia patient education tool: a pilot study. *Foot Ankle Int* 2009 Jul 01;30(7):619-626. [doi: [10.3113/fai.2009.0619](https://doi.org/10.3113/fai.2009.0619)]
4. Yin B, Goldsmith L, Gambardella R. Web-based education prior to knee arthroscopy enhances informed consent and patient knowledge recall: a prospective, randomized controlled study. *J Bone Joint Surg Am* 2015 Jun 17;97(12):964-971. [doi: [10.2106/JBJS.N.01174](https://doi.org/10.2106/JBJS.N.01174)] [Medline: [26085529](https://pubmed.ncbi.nlm.nih.gov/26085529/)]
5. Cornoiu A, Beischer AD, Donnan L, Graves S, de Steiger R. Multimedia patient education to assist the informed consent process for knee arthroscopy. *ANZ J Surg* 2011 Mar;81(3):176-180. [doi: [10.1111/j.1445-2197.2010.05487.x](https://doi.org/10.1111/j.1445-2197.2010.05487.x)] [Medline: [21342392](https://pubmed.ncbi.nlm.nih.gov/21342392/)]
6. Jjala HA, French JL, Foxall GL, Hardman JG, Bedfordth NM. Effect of preoperative multimedia information on perioperative anxiety in patients undergoing procedures under regional anaesthesia. *Br J Anaesth* 2010 Mar;104(3):369-374 [FREE Full text] [doi: [10.1093/bja/aeq002](https://doi.org/10.1093/bja/aeq002)] [Medline: [20124283](https://pubmed.ncbi.nlm.nih.gov/20124283/)]
7. Migden M, Chavez-Frazier A, Nguyen T. The use of high definition video modules for delivery of informed consent and wound care education in the Mohs Surgery Unit. *Semin Cutan Med Surg* 2008 Mar;27(1):89-93. [doi: [10.1016/j.sder.2008.02.001](https://doi.org/10.1016/j.sder.2008.02.001)] [Medline: [18486030](https://pubmed.ncbi.nlm.nih.gov/18486030/)]
8. Beischer AD, Clarke A, de Steiger RN, Donnan L, Ibuki A, Unglik R. The practical application of multimedia technology to facilitate the education and treatment of patients with plantar fasciitis: a pilot study. *Foot Ankle Spec* 2008 Feb;1(1):30-38. [Medline: [19825689](https://pubmed.ncbi.nlm.nih.gov/19825689/)]
9. Maasland E, Koudstaal PJ, Habbema JD, Dippel DW. Effects of an individualized multimedia computer program for health education in patients with a recent minor stroke or transient ischemic attack - a randomized controlled trial. *Acta Neurol Scand* 2007 Jan;115(1):41-48. [doi: [10.1111/j.1600-0404.2006.00722.x](https://doi.org/10.1111/j.1600-0404.2006.00722.x)] [Medline: [17156264](https://pubmed.ncbi.nlm.nih.gov/17156264/)]
10. Kandula NR, Malli T, Zei CP, Larsen E, Baker DW. Literacy and retention of information after a multimedia diabetes education program and teach-back. *J Health Commun* 2011;16 Suppl 3:89-102. [doi: [10.1080/10810730.2011.604382](https://doi.org/10.1080/10810730.2011.604382)] [Medline: [21951245](https://pubmed.ncbi.nlm.nih.gov/21951245/)]
11. Chen H, Yeh M, Yang H. Testing the impact of a multimedia video CD of patient-controlled analgesia on pain knowledge and pain relief in patients receiving surgery. *Int J Med Inform* 2005 Jul;74(6):437-445. [doi: [10.1016/j.ijmedinf.2005.04.003](https://doi.org/10.1016/j.ijmedinf.2005.04.003)] [Medline: [15936246](https://pubmed.ncbi.nlm.nih.gov/15936246/)]
12. Huber J, Ihrig A, Yass M, Bruckner T, Peters T, Huber CG, et al. Multimedia support for improving preoperative patient education: a randomized controlled trial using the example of radical prostatectomy. *Ann Surg Oncol* 2013 Jan;20(1):15-23. [doi: [10.1245/s10434-012-2536-7](https://doi.org/10.1245/s10434-012-2536-7)] [Medline: [22851045](https://pubmed.ncbi.nlm.nih.gov/22851045/)]
13. McNicol L, Hutchinson A, Wood B, Botti M, Redley B. An exploration of patients' experience of nurses' use of point-of-care information technology in acute care. *Patient Experience J* 2018 Apr 25;5(1):76-89 [FREE Full text] [doi: [10.35680/2372-0247.1213](https://doi.org/10.35680/2372-0247.1213)]
14. Brown D, McCormack B. Developing postoperative pain management: utilising the promoting action on research implementation in health services (PARIHS) framework. *Worldviews Evid Based Nurs* 2005;2(3):131-141. [doi: [10.1111/j.1741-6787.2005.00024.x](https://doi.org/10.1111/j.1741-6787.2005.00024.x)] [Medline: [17040534](https://pubmed.ncbi.nlm.nih.gov/17040534/)]
15. Attitudes and perceptions of advanced practice nurses towards health information technology and its effects on caring. San Jose State University. URL: https://scholarworks.sjsu.edu/etd_doctoral/59/ [accessed 2022-05-04]

16. McDonall J, de Steiger R, Reynolds J, Redley B, Livingston PM, Hutchinson AF, et al. Patient activation intervention to facilitate participation in recovery after total knee replacement (MIME): a cluster randomised cross-over trial. *BMJ Qual Saf* 2019 Oct;28(10):782-792 [[FREE Full text](#)] [doi: [10.1136/bmjqs-2018-008975](https://doi.org/10.1136/bmjqs-2018-008975)] [Medline: [30975730](https://pubmed.ncbi.nlm.nih.gov/30975730/)]
17. O'Leary KJ, Lohman ME, Culver E, Killarney A, Randy Smith G, Liebovitz DM. The effect of tablet computers with a mobile patient portal application on hospitalized patients' knowledge and activation. *J Am Med Inform Assoc* 2016 Jan;23(1):159-165 [[FREE Full text](#)] [doi: [10.1093/jamia/ocv058](https://doi.org/10.1093/jamia/ocv058)] [Medline: [26078412](https://pubmed.ncbi.nlm.nih.gov/26078412/)]
18. Craig P, Dieppe P, Macintyre S, Michie S, Nazareth I, Petticrew M. Developing and evaluating complex interventions: the new Medical Research Council guidance. *Int J Nurs Stud* 2013 May;50(5):587-592. [doi: [10.1016/j.ijnurstu.2012.09.010](https://doi.org/10.1016/j.ijnurstu.2012.09.010)] [Medline: [23159157](https://pubmed.ncbi.nlm.nih.gov/23159157/)]
19. Grimshaw J, Eccles M, Lavis J, Hill SJ, Squires JE. Knowledge translation of research findings. *Implement Sci* 2012 May 31;7:50 [[FREE Full text](#)] [doi: [10.1186/1748-5908-7-50](https://doi.org/10.1186/1748-5908-7-50)] [Medline: [22651257](https://pubmed.ncbi.nlm.nih.gov/22651257/)]
20. Atkins L, Francis J, Islam R, O'Connor D, Patey A, Ivers N, et al. A guide to using the Theoretical Domains Framework of behaviour change to investigate implementation problems. *Implement Sci* 2017 Jun 21;12(1):77 [[FREE Full text](#)] [doi: [10.1186/s13012-017-0605-9](https://doi.org/10.1186/s13012-017-0605-9)] [Medline: [28637486](https://pubmed.ncbi.nlm.nih.gov/28637486/)]
21. Bolster D, Manias E. Person-centred interactions between nurses and patients during medication activities in an acute hospital setting: qualitative observation and interview study. *Int J Nurs Stud* 2010 Feb;47(2):154-165. [doi: [10.1016/j.ijnurstu.2009.05.021](https://doi.org/10.1016/j.ijnurstu.2009.05.021)] [Medline: [19577752](https://pubmed.ncbi.nlm.nih.gov/19577752/)]
22. Jangland E, Carlsson M, Lundgren E, Gunningberg L. The impact of an intervention to improve patient participation in a surgical care unit: a quasi-experimental study. *Int J Nurs Stud* 2012 May;49(5):528-538. [doi: [10.1016/j.ijnurstu.2011.10.024](https://doi.org/10.1016/j.ijnurstu.2011.10.024)] [Medline: [22104043](https://pubmed.ncbi.nlm.nih.gov/22104043/)]
23. Taylor J, Rutherford P. The pursuit of genuine partnerships with patients and family members: the challenge and opportunity for executive leaders. *Front Health Serv Manage* 2010;26(4):3-14. [Medline: [20617664](https://pubmed.ncbi.nlm.nih.gov/20617664/)]
24. McTier L, Botti M, Duke M. Patient participation in medication safety during an acute care admission. *Health Expect* 2015 Oct 17;18(5):1744-1756 [[FREE Full text](#)] [doi: [10.1111/hex.12167](https://doi.org/10.1111/hex.12167)] [Medline: [24341439](https://pubmed.ncbi.nlm.nih.gov/24341439/)]
25. McTier L, Botti M, Duke M. Patient participation in pulmonary interventions to reduce postoperative pulmonary complications following cardiac surgery. *Aust Crit Care* 2016 Feb;29(1):35-40. [doi: [10.1016/j.aucc.2015.04.001](https://doi.org/10.1016/j.aucc.2015.04.001)] [Medline: [25939547](https://pubmed.ncbi.nlm.nih.gov/25939547/)]
26. McTier LJ, Botti M, Duke M. Patient participation in quality pain management during an acute care admission. *Clin J Pain* 2014 Apr;30(4):316-323. [doi: [10.1097/AJP.0b013e318298ddc5](https://doi.org/10.1097/AJP.0b013e318298ddc5)] [Medline: [23727655](https://pubmed.ncbi.nlm.nih.gov/23727655/)]
27. Westbrook JI, Duffield C, Li L, Creswick NJ. How much time do nurses have for patients? A longitudinal study quantifying hospital nurses' patterns of task time distribution and interactions with health professionals. *BMC Health Serv Res* 2011 Nov 24;11:319 [[FREE Full text](#)] [doi: [10.1186/1472-6963-11-319](https://doi.org/10.1186/1472-6963-11-319)] [Medline: [22111656](https://pubmed.ncbi.nlm.nih.gov/22111656/)]
28. Blackman I, Henderson J, Willis E, Hamilton P, Toffoli L, Verrall C, et al. Factors influencing why nursing care is missed. *J Clin Nurs* 2015 Jan;24(1-2):47-56. [doi: [10.1111/jocn.12688](https://doi.org/10.1111/jocn.12688)] [Medline: [25265893](https://pubmed.ncbi.nlm.nih.gov/25265893/)]
29. Cook DJ, Moradkhani A, Douglas KS, Prinsen SK, Fischer EN, Schroeder DR. Patient education self-management during surgical recovery: combining mobile (iPad) and a content management system. *Telemed J E Health* 2014 Apr;20(4):312-317 [[FREE Full text](#)] [doi: [10.1089/tmj.2013.0219](https://doi.org/10.1089/tmj.2013.0219)] [Medline: [24443928](https://pubmed.ncbi.nlm.nih.gov/24443928/)]
30. Schnock KO, Snyder JE, Fuller TE, Duckworth M, Grant M, Yoon C, et al. Acute care patient portal intervention: portal use and patient activation. *J Med Internet Res* 2019 Jul 18;21(7):e13336 [[FREE Full text](#)] [doi: [10.2196/13336](https://doi.org/10.2196/13336)] [Medline: [31322123](https://pubmed.ncbi.nlm.nih.gov/31322123/)]
31. Cook DJ, Manning DM, Holland DE, Prinsen SK, Rudzik SD, Roger VL, et al. Patient engagement and reported outcomes in surgical recovery: effectiveness of an e-health platform. *J Am Coll Surg* 2013 Oct;217(4):648-655. [doi: [10.1016/j.jamcollsurg.2013.05.003](https://doi.org/10.1016/j.jamcollsurg.2013.05.003)] [Medline: [23891066](https://pubmed.ncbi.nlm.nih.gov/23891066/)]
32. Chu ES, Hakkarinen D, Evig C, Page S, Keniston A, Dickinson M, et al. Underutilized time for health education of hospitalized patients. *J Hosp Med* 2008 May;3(3):238-246. [doi: [10.1002/jhm.295](https://doi.org/10.1002/jhm.295)] [Medline: [18570334](https://pubmed.ncbi.nlm.nih.gov/18570334/)]
33. Dillon A. User acceptance of information technology. In: *Encyclopedia of Human Factors and Ergonomics*. London: Taylor and Francis; 2001.
34. Greysen SR, Magan Y, Rosenthal J, Jacobia R, Auerbach AD, Harrison JD. Patient recommendations to improve the implementation of and engagement with portals in acute care: hospital-based qualitative study. *J Med Internet Res* 2020 Jan 14;22(1):e13337 [[FREE Full text](#)] [doi: [10.2196/13337](https://doi.org/10.2196/13337)] [Medline: [31934868](https://pubmed.ncbi.nlm.nih.gov/31934868/)]
35. Vardoulakis L, Karlson A, Morris D, Smith G, Gatewood J, Tan D. Using mobile phones to present medical information to hospital patients. In: *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. 2012 Presented at: CHI '12: CHI Conference on Human Factors in Computing Systems; May 5 - 10, 2012; Austin Texas USA. [doi: [10.1145/2207676.2208601](https://doi.org/10.1145/2207676.2208601)]

Abbreviations

MyStay TKR: MyStay Total Knee Replacement

Edited by A Kushniruk; submitted 31.01.22; peer-reviewed by A van der Horst, J Sarwark; comments to author 26.03.22; revised version received 04.04.22; accepted 23.04.22; published 19.05.22.

Please cite as:

McDonall J, Redley B, Livingston P, Hutchinson A, de Steiger R, Botti M

A Nurse-Led Multimedia Intervention to Increase Patient Participation in Recovery After Knee Arthroplasty: Hybrid Type II Implementation Study

JMIR Hum Factors 2022;9(2):e36959

URL: <https://humanfactors.jmir.org/2022/2/e36959>

doi: [10.2196/36959](https://doi.org/10.2196/36959)

PMID: [35588363](https://pubmed.ncbi.nlm.nih.gov/35588363/)

©Jo McDonall, Bernice Redley, Patricia Livingston, Ana Hutchinson, Richard de Steiger, Mari Botti. Originally published in JMIR Human Factors (<https://humanfactors.jmir.org>), 19.05.2022. This is an open-access article distributed under the terms of the Creative Commons Attribution License (<https://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Human Factors, is properly cited. The complete bibliographic information, a link to the original publication on <https://humanfactors.jmir.org>, as well as this copyright and license information must be included.

Original Paper

Co-creation of a Serious Game About Radiotherapy: Participatory Action Research Study With Children Treated for Cancer

Catarina Cederved¹, MScN; Jon Back², PhD; Charlotte Ångström-Brännström^{1,3}, PhD; Gustaf Ljungman¹, MD, PhD; Gunn Engvall¹, PhD

¹Department of Women's and Children's Health, Uppsala University, Uppsala, Sweden

²Department of Informatics and Media, Uppsala University, Uppsala, Sweden

³Department of Nursing, Umeå University, Umeå, Sweden

Corresponding Author:

Catarina Cederved, MScN

Department of Women's and Children's Health

Uppsala University

Akademiska sjukhuset

Uppsala, 75185

Sweden

Phone: 46 734697207

Email: catarina.cederved@kbh.uu.se

Abstract

Background: Children with cancer who have to undergo radiotherapy can experience fear, because they have no prior knowledge of the treatment. One way of teaching children about the treatment and reducing their fear is to prepare them for it through serious games. Involvement of the end user in the design process within medicine is a way of ensuring that the product being developed will fit the intended user.

Objective: The aim was to outline the contributions made by children and their parents through participatory action research when designing a serious game about radiotherapy.

Methods: By means of participatory action research, children and their parents participated in the development of a serious game about radiotherapy. Nine children (7-10 years old) were included, each with an accompanying parent. A qualitative approach was used that included interviews and participant observation. Six rounds of iterative development process were used with the children and their parents. Meetings with the children were held either face-to-face or online. Each round resulted in a list of suggestions for changes to the game. A thematic analysis was performed based on the list of proposed changes, underpinned by all gathered data, to highlight how the children's participation changed the game.

Results: Two main themes were identified. The first theme was "The children's participation was affected by their health and treatment" and included the following subthemes: "an opportunity to share emotions and perceptions of radiotherapy" and "the possibility to participate was affected by the severity of the disease." The second theme was "participation allowed becoming an active part of game development" and included the following subthemes: "the opportunity to express sentiments about the game," "the emergence of a playable game through the children's contributions," and "the necessity of understanding the text."

Conclusions: The method used in this study made the children active participants, and our results suggest that this method can be used by health care researchers to cocreate serious games with children. It is necessary to inform the children involved that the process takes time, and that the process can be altered to allow as much participation as possible without placing a burden on them. The children's illness affected their possibility to take part; thus, it is crucial to accommodate the children's needs when conducting similar studies. The parents' participation facilitated the meetings for their children, even though their involvement in the game design was negligible.

(*JMIR Hum Factors* 2022;9(2):e34476) doi:[10.2196/34476](https://doi.org/10.2196/34476)

KEYWORDS

children; participatory action research; game design; radiotherapy; education; supportive care; oncology

Introduction

Children with cancer face many challenges and distressing events [1]. The procedures that the children have to undergo when receiving treatment are previously unknown to them and often cause fear [2,3]. Fear of the unknown has been described as a fundamental fear, and in situations with unknown elements, patients should be provided with tools to increase their ability to cope [4]. Radiotherapy (RT) is one of the major treatment modalities that can cure or alleviate cancer, depending on the cancer diagnosis [5]. For children, it can be difficult to receive RT since they must remain in a fixed position without moving, and they are left in the treatment room by themselves [6]. Some children need daily anesthesia to cope with the procedure, lasting for several weeks and accompanied by fasting periods; this is suboptimal for a growing body [6]. It has been suggested that with the right preparation children may have a greater chance of coping with the RT procedure, allowing them to receive the treatment without sedation [7-9].

The current project is based on previous research conducted by Engvall et al [10], who developed a digital story describing the RT procedure that was delivered as an application on a tablet designed for children. The participating children suggested that the application could have been more interactive if it had been designed as a game [10]. When designing serious games such as these, there are several aspects that need to be considered, such as the purpose of the game, the end users, the stakeholders in the project, and how the game can engage the players [11-13]. Serious games are a way to learn and are meant to solve real world problems; in the best case, they are enjoyable and entertaining [14,15]. Involving the end user in development has become common practice when developing games and interactive applications for children in medical facilities [16,17], as this is considered to increase the chances of creating a successful game [18]. Participatory action research (PAR) involves children as cocreators in the research process, allowing them to contribute and take part at different levels of the study [19]. Thus, PAR can strengthen study findings and the interventions being researched [20]. To our knowledge, no serious games have yet been developed to prepare children for the RT procedure. Further, researchers of serious game design have highlighted the significance of describing and reporting the involvement of end users in the design process and in research during the design process [21]. This paper is not primarily meant to point out the relevance of participatory design. Rather, it is meant to show how PAR can be used to bring about changes, using RT as an example. We wanted to create a serious game about RT in collaboration with children with experience of RT, to promote knowledge of the procedure in children who will receive RT in the future. The aim was to outline the contributions made by children and their parents through PAR when designing a serious game about RT.

Methods

Study Design

The present study describes the development of a serious game for RT as part of a larger project to reduce RT-related anxiety.

The development of the game used PAR by involving children that had undergone RT, their parents, an expert team of health care professionals, game designers, and a research team. A qualitative approach was used that included interviews and participant observation.

Recruitment of Participants

A nurse at a pediatric oncology center contacted parents with a child who had received or was receiving RT to introduce the study. The inclusion criteria were that the child had received or was receiving RT, was fluent in Swedish, and was between 6 and 15 years old. A total of 14 families were approached, of whom 5 declined to participate. Explanations given for declining included lack of interest in a game project, not having time, and prioritizing school; some families gave no explanation.

Ethical Approval

Informed and written consent was given by parents on their child's behalf and by the parents on their own behalf. Written assent was given by the children participating in the study. The study design was approved by the Regional Ethical Review Board in Uppsala, Sweden (2018/264).

Data Collection

Before each meeting, the parents were contacted with appointment suggestions. Meetings were scheduled at times that most of the children could attend and were held in 2 towns to shorten the traveling distance for the families, meaning the participants were divided into 2 groups. The families were reimbursed for their travel expenses.

The game development was performed as an iterative process [22]. The children met at play meetings where they played and commented on a prototype of the game. The children were observed while playing the game prototype and their gameplay was recorded on a computer. The children were asked questions about elements in the game to further understand how they interpreted them. After the game testing, a semistructured interview with follow-up questions was conducted in a group or individually; these interviews were audiotaped (an interview guide is included in [Multimedia Appendix 1](#)).

In the first round, the children and accompanying parents met at the research venue and were offered refreshments. The children discussed their experiences of RT, what games they preferred playing, and talked freely about what they thought should be in the game that was going to be developed. Two investigators were present; one was active and interacted with the children, while the other observed and took notes. In an effort to make the children more comfortable and not feel inferior in the situation, the decision was made to limit the number of adults by not having the design team present. This also solved a logistical problem: the design team was located in a city several hours away from where the study was conducted. In the second round, the children played a prototype of the game in groups of 2 or individually. After the second iteration, a change was made due to COVID-19 to minimize potential infection. The families were given options to meet with one investigator and one other family, meet with only the investigator, or to have a virtual meeting with one investigator

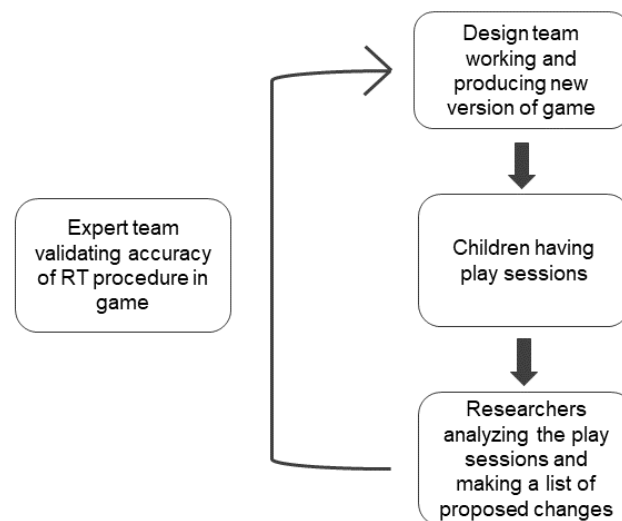
over the internet. However, the rounds still had the same structure: the children played the game prototype, sometimes showed their parents the game, and talked about it with their parents if they wanted to. The parents were asked in the later rounds to provide their thoughts and experiences of the game.

Analysis and Game Design Process

The material analyzed from each round with the children and their accompanying parents consisted of 1) screen recordings of gameplay, 2) observational field notes taken during gameplay, 3) summary notes made after each round, and 4) audio recordings of the subsequent interviews. The screen recordings were viewed multiple times; inductive coding was performed immediately after each round. In this way, 3 main coding categories emerged. These were audiovisual cues, game mechanics, and narrative. Therefore, coding of the early rounds

informed later rounds. The material from later rounds was coded into the discovered categories. A similar approach was then used to identify codes for the field notes, summary notes, and interviews. All 4 data sets were compared to find commonalities between the children's play. Codes were also grouped, revealing design suggestions. The suggestions were then prioritized into a list of proposed changes that was presented to the design team within 1 to 2 weeks after each round. After discussions on feasibility and time, the design team made as many of the changes as possible within the given timeframe. The new changes to the game prototype were observed at the next iteration (Figure 1). The iterative procedure with the children was repeated 6 times over a period of 8 months during 2020. None of the children took part in the analysis or the processing of the material.

Figure 1. The iteration process. RT: radiotherapy.



A thematic analysis based on the content from the list of proposed changes underpinned by all the gathered data was later performed to capture the children's and parents' participation in changes made to the game. Data were analyzed through a thematic analysis inspired by Braun and Clark [23]. Codes were identified and grouped and themes and subthemes were formulated, discussed with coauthors, and revised when necessary until consensus was obtained. Finally, quotations and figures were added to illustrate the content.

Description of the Game

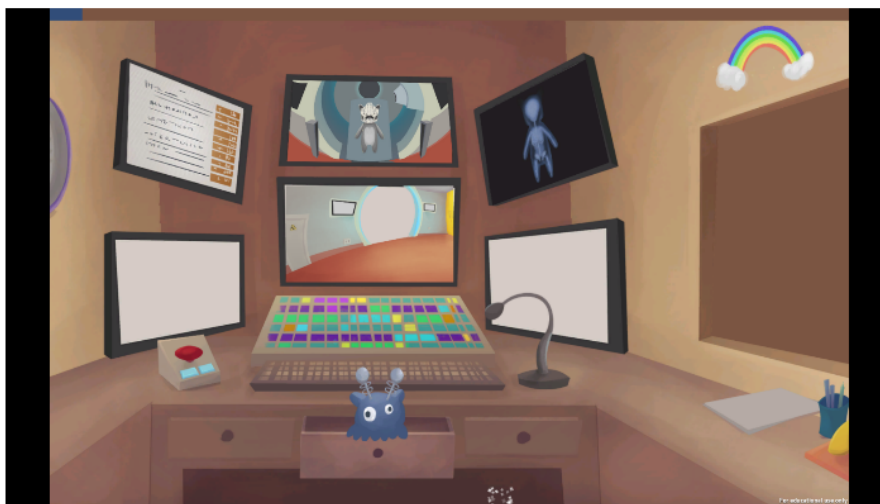
The design team that was assembled to build the game consisted of students from the Department of Game Design at the University of Uppsala and a lecturer who provided supervision. The first version of the game prototype came about after a workshop with the expert team, research team, and design team identified problems, chose the technological platform, and identified the end users. The workshop was held at the clinic that was later portrayed in the game. The first version was a proof of concept developed for children 9 to 15 years old. The game used a third person perspective and was set in a cartoon-style RT clinic where the player followed an avatar through the RT process. The game consisted of linear puzzle

rooms in which the player had to complete certain tasks in order to unlock subsequent rooms and tasks. The first version included rooms depicting an RT room, a monitoring room, and a bedroom, as well as mini games not related to the main puzzle. The game was played as a linear loop. The final version of the game came about after the sixth round with the participants and included rooms depicting a reception area, an RT room, a monitoring room (Figure 2 and Figure 3), a narcosis room, a practice room for RT, a kitchen, and a bedroom, which included hidden scenes of a wardrobe and an outdoor environment. The need to complete tasks to be able to progress within the game had been removed, and the game was now played as a "doll-house" style game [24]. The rooms were connected through a game map and all were already accessible from the beginning of play. Every room or environment depicted in the game actually existed at the hospital, apart from the practice room. Some of the mini games corresponded to the RT procedure, while others did not. Some components not relevant to the RT procedure were included in the game to encourage the player to play for a longer time, to encourage engagement with the RT scenes, and to promote understanding of the RT procedure.

Figure 2. Radiotherapy room after changes were made based on the children's and parents' contributions.



Figure 3. Monitoring room after changes were made based on the children's and parents' contributions.



Results

Description of the Data

Iteration of the game design took place over 6 rounds, each of which included testing with the children, and each of which resulted in a new version of the prototype. Each round included multiple meetings held at different times depending on when the children could attend. The study included 8 children in the first round, and later, 1 additional participant was recruited. [Table 1](#) shows information on how many children participated in each round. All the children had a single parent with them

during each round, except for 1 child whose parents took turns accompanying the child. At the time they participated, the children were between 7 and 10 years old and included 1 boy and 8 girls. No child was 9 years old at the start of the project. The children had various diagnoses, including brain tumor, rhabdomyosarcoma, spinal cord cancer, and Ewing sarcoma. Two children chose to leave the study in the fourth round.

Two main themes were revealed by the analysis: (1) the children's participation was affected by their health and treatment, and (2) participation allowed becoming an active part of game development. Each theme had subthemes that described aspects of the cocreative process.

Table 1. Number of participating children in each round and meeting (N=9).

Number of participating children in each meeting, n	Round number					
	Round 1	Round 2	Round 3	Round 4 ^a	Round 5	Round 6
Meeting 1	3	4	2	2	1	1
Meeting 2	4	1	2	2	1	1
Meeting 3	0	0	1	1	1 ^b	1
Meeting 4	0	0	1 ^b	0	1	1 ^b
Meeting 5	0	0	1 ^b	0	0	1 ^b
Meeting 6	0	0	1 ^b	0	0	0
Total	7	5	8	5	4	5

^aRound included newly recruited participant.

^bRound/meeting was conducted online.

The Children's Participation Was Affected by Their Health and Treatment

The children and parents shared how they felt and what it was like when they received RT. Their stories became subthemes: (1) how the meetings provided an opportunity to share emotions and perceptions of RT, and (2) how the possibility to participate was affected by the severity of the disease.

An Opportunity to Share Emotions and Perceptions of RT

The children described by means of emotion cards how they felt about receiving RT. When talking about RT, a few said that they had been happy, but they did not know how it worked when they received their first treatment. However, most of the children said that they had been worried before receiving RT, because they were not sure what would happen or how it would feel (Figure 4). Hence, the game needed to capture the essence of the RT procedure from a child's point of view.

Figure 4. A child's chosen emotion cards for the emotions the child felt during the first radiotherapy treatment. The cards were printed with the approval of St Luke's Innovative Resources.



One parent expressed that they had felt that their child looked vulnerable due to the number of staff around the child when preparing for RT. After checking with the expert team regarding how many people were usually present, additional personnel were added in the game's RT room in accordance with the input from the parents.

The Possibility to Participate Was Affected by the Severity of the Disease

Due to the children's illness, some of them showed signs of fatigue or lack of strength. Some children had to undergo treatment at the hospital in the morning before coming to a meeting, and some children had to choose between going to school or participating in the study, because they did not have the strength to do both. If they found the testing interesting,

they stayed, even though they sometimes showed tiredness. The audio recording of one of the meetings in round 1 with 4 children has audible yawning more than once toward the end. The meeting lasted for a little over an hour. Therefore, the day and length of the meetings were modified and adapted to meet the children's ability to participate. The families could choose to have the meetings online or could choose a date within a certain week that worked for them.

Participation Allowed Becoming an Active Part of Game Development

Three subthemes were identified within the main theme: (1) the opportunity to express sentiments about the game; (2) the emergence of a playable game through the children's contributions; and (3) the necessity of understanding the text. Examples are given below of how the children actively contributed to the game design process.

The Opportunity to Express Sentiments About the Game

During the first meeting, the children discussed other games they liked to play and explained features they thought should be developed in our game. They reported contrasting gaming habits. Some were used to playing very advanced video games while others hardly played at all. The children agreed that the main purpose of the game was to explain how RT worked, to follow the character through the procedure, and to give the character medication. However, they added that the game should be fun and exciting. Their input led to the first prototype of the game being scrapped due to its long introduction. The game was reconstructed to start with RT. After playing the game, the 10-year-old children and their parents thought that the appearance of the game was similar to the real RT setting, but that the game was boring and childish with too little explanation of how the RT procedure worked. One child said that they wanted a scene that showed what happened when you were sedated for RT. In response to these comments, more information was added to the RT room and an extra scene was added. A narcosis room was also added, as well as a practice room where the player could encounter different elements of RT through comic strips. However, we did not obtain ideas on how to make the game more interesting or amusing. Some of the children who were 7 to 8 years old were under anesthesia during RT, so they could not comment on the accuracy of the portrayal, although they enjoyed the interactions that were in the game for amusement and creative play, especially the mini games. One child said they really enjoyed the game.

The children were shown a map of the game on paper and asked what they thought about it. They gave no opinion as to whether they considered the map good or bad, but when it was implemented in the game, thereby removing the game's linearity, the children spontaneously used it and understood its function. One of the 10-year-old children was asked in the last round why they returned to the kitchen in the game, even though nothing new had been implemented there, and they answered: "Because it is fun there." When asked why, they replied, "Because I can pick tomatoes." Another child, in the 7- to 8-year-old age group, also stated that they thought the game was fun to play because of particular scenes and interactivity.

The same child had earlier in the iteration process exclaimed that they thought the game was boring and stupid.

A major feature of the game was teaching the player strategies on how to cope with the RT procedure. However, in the interviews, the children pointed out that one of these coping strategies was incorrectly portrayed. Children that undergo RT are sometimes given a string to hold on to; a parent, in an adjacent room, holds the other end and can pull the string so that the child feels the tug. In the game, the child character was shown holding onto a string, and to visualize the tug, the character's hand was shown moving. However, the children explained that in fact, they were not allowed to move at all during RT, so the mini game was removed and replaced with a comic strip explaining the strategy.

When discussing what the children did to keep themselves occupied during RT, they said that they often thought about things they would do later that day. One child explained that they thought about food they would like to eat. These accounts from the children led to their specific coping strategies being added to the game's RT room.

The Emergence of a Playable Game Through the Children's Contributions

The children sometimes had trouble with movement in the game, while at other times, they misunderstood the game environment and missed gameplay elements that were intended to give clues on how the game worked. We therefore added a visual aid to indicate the location of items and how to advance the story. These additions allowed the children to play autonomously. Some children had problems using a mouse to control the game, which sometimes irritated them, since it meant that they could not, for example, pick up an item. To facilitate their understanding of how to play and interact with the game, we added an introductory scene as a tutorial, set in a reception area. After reading the tutorial, the children said they understood how to maneuver through the game, but when we asked them how to do it, we realized that they still had some problems. The children's feedback on the tutorial scene was that it looked like the clinic. Some children asked what they were meant to do when introduced to new scenes during play. The game already featured some scenes with telephones that informed the player what they could do in that scene, so we added them to every scene. Some children found that the sound effects in the game were too loud, especially when playing online with headphones. Therefore, the sound effects were modified and a function was added that let the player turn the music and sound effects off and on and read the text in the game aloud.

Some elements of linearity were kept in the game. For example, the RT procedure in the RT room ran on a loop. The parents questioned this, arguing that the player should not have to sit through the loop before leaving the room. Furthermore, one child in the recordings could be heard saying, "Okay, I'll play through this part," only because they wanted to get through it to reach the next part of the game. Hence, the game was changed to let the player exit the RT room via the map, giving them the choice to play the loop or skip it.

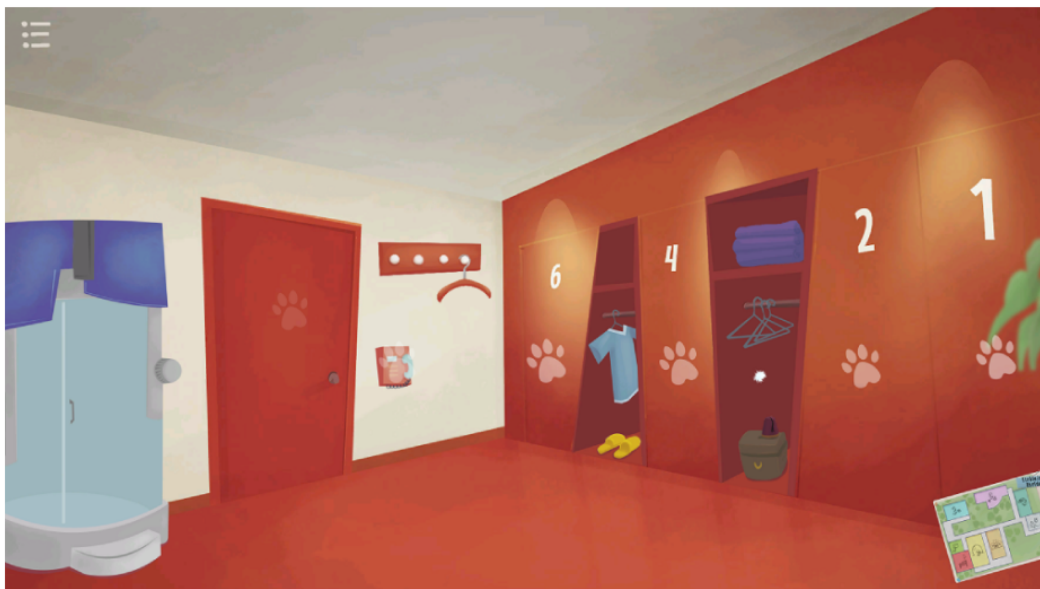
A kitchen scene had a tap that started to dance and make elephant noises when the player clicked it. Two of the 7- to 8-year-old children, who were playing together, clicked on the tap several times and talked about it. One of them showed it to their parent and said, “You need to wash your hands,” and then laughed happily and called it funny. Even the older 10-year-old children laughed at the tap. Inspired by this, we added back, in random, hidden places, scenes that had previously been deleted.

When the children encountered these hidden scenes, they showed them either to each other or to their parents and talked about them. When one child played the game, they commented, “This room is so cool!” explaining, “But the closet is so small, and then all of this shows up.” The interaction consisted of a changing room that was hidden within a closet (Figure 5 and Figure 6). Children who found these hidden scenes started to search other scenes more carefully.

Figure 5. Clicking on the paw symbol on the closet reveals a hidden scene.



Figure 6. A hidden scene (a walk-in closet) revealed by the interaction in Figure 5.



A practice room was added to give more information about RT. One of the 7- to 8-year-old children showed particular interest in a prototype of a timer that allowed visualizing how to be still for a length of time. When asked what they were doing, the child replied, “I am practicing to see how long I can be still.”

This apparent interest in the timer inspired us to portray more elements of RT in a way that would create value. This led us to develop the idea of using comic strips to explain certain elements of the RT procedure (Figure 7).

Figure 7. Comic strip explaining that it is normal to feel scared of the procedure, and that this feeling usually ceases after becoming accustomed to it; the monsters under the bed are an analogy.



In the last round, the children showed awareness that the game had changed extensively from the first time they played it and agreed that it gave a valid introduction to RT while including elements for amusement. In all, numerous changes had been made to the game due to the children's participation. Examples of some, but not all, of these changes are included in [Multimedia Appendix 2](#).

The Necessity of Understanding the Text

A significant number of the learning goals in the game were delivered through text snippets, and understanding these snippets was consequently an integral part of the player experience. Some children struggled with the amount of text; this proved most difficult for the 7- to 8-year-old children, due to their lack of reading skills. When a researcher offered to read the text, almost all the children in this age group preferred this. Reading the text gave us instant feedback about how well it worked, and allowed us to revise the text with every new version to better accommodate the 7- to 8-year-old players. The game's final version included audio of the text that the player could activate.

We also implemented comic strips with explanatory text to increase understanding of what RT is like and what can happen when receiving RT. All these comics used metaphors for various elements of the treatment ([Figure 7](#)). The comics also used relatable, real-life analogies to describe and create an understanding of what could happen when receiving RT.

The children that read the comic strips were asked about them. Some children asked questions while reading them. These questions were often about the text; the children were sometimes confused when the text did not correspond to the picture or when the same text was used for a different picture. However, when they read the entire comic strip, they seemed to understand its overall purpose. For example, the comics used a metaphor of "monsters under the bed" that the children related to and thought could be added to the game ([Figure 7](#)).

Discussion

Summary

Our findings represent a thorough description of how children were part of the development of a serious game aimed at increasing knowledge about RT and decreasing anxiety related to the procedure. We performed a thematic analysis that revealed two main themes and several subthemes. The main themes were that the children's participation was affected by their health and treatment, and that participation allowed becoming an active part of game development. The children's participation influenced development of the game and led to changes to the

game's narrative, mechanics, and aesthetics. The game became more of a "doll-house" style experience, focusing on exploration and discovery, with the addition of a map that made it possible to move freely between the rooms [25]. With the exception of mini games, winning conditions were removed to make space for free, unstructured play. To ensure that the younger, 7- to 8-year-old children could manage to play, their input on game mechanics was prioritized, because they are more often in need of anesthesia and therefore more likely to be the end users. The participation of the 10-year-old children added knowledge about the RT procedure and the depiction of the game environment. The parents' participation did not lead to more than a few changes, but their participation facilitated the meetings for the children.

Principal Findings of Participation

Severe Illness Affects the Possibility to Participate

It proved to be crucial to modify the meetings to accommodate the children's need for rest due to their illness and treatment, so that they could continue their participation. Discussing a meeting's time and place with the parents and the participants prior to the meeting was a way to adjust the requirements. To include participants in the decisions on when, where, and for how long to hold the meetings enabled the participants to be part of the development of the game, as well as the design of the study [26]. This change was primarily made due to COVID-19, rather than the children's illnesses as such, but it nevertheless facilitated participation. This finding is imperative for researchers planning to involve similar children. Children are considered to be a vulnerable group due to their stage of cognitive development; they are often regarded as not being able to fully comprehend the research in which they are taking part. The group in this study was also considered vulnerable due to the severity of their illnesses [27]. Nevertheless, it has been established that including children in the development of future interventions is necessary, because the intervention concerns them as a group [26]. However, since our group could not meet all at once, the children themselves could not reach a consensus on what changes were needed in the game; that had to be established by the lead investigator. Hence, the investigator had to search the children's game play and interviews for commonalities and differences in the material and then create a list of proposed changes to present to the design team. This is not a common way of working in developmental design, but it might be a plausible way to enable seriously ill 7- to 11-year-old children to be part of the process. Further, when using PAR, children are not only a source of data but a part of development [28] and are partners capable of contributing to

the results [29]. The children who participated in this study became both stakeholders in the product under development and also the voices of the end users. On Shier's ladder of participation [30], the children were somewhere between "shared decisions with children" and "child initiated and directed participation." While the design decisions were made by the research team and game designers, these decisions were highly informed by the children's and parents' input.

Children's Contributions Led to More Comprehensible Gameplay

The game needed to be comprehensible to the children, both when it came to how and why to play and also the informational content; the game's interactivity also relied on knowing how to control it.

The children's constant feedback on how they thought the game worked allowed us to design the game to be suitable for them. As an example, the children influenced how the coping strategies were presented in the game. They commented on the designer's mistakes, such as the previously mentioned visualization of hand movement when a parent tugged on a string, when in reality they had to be completely still, and pointed out other coping strategies they had used that were not displayed in the game. The children's involvement in the design process allowed us to correct mistakes and add new coping strategies to make the game more accurate and provide end users more tools. Exaggerating motions, such as the hand moving to represent the tug of the string, is a common way of presenting this type of movement in animation [31]. However, as the children experienced, and pointed out to us, it was important to not move throughout the treatment, so this exaggeration was something they understood as a problem. As they had personal experience of RT, this issue was visible to them in a way that would be very hard for anyone else to see. While it may seem like an obvious thing, it was not obvious until they pointed it out. Thus, their expertise in being children gave the designers a unique understanding of the situation. With the children's participation, the designers ensured that the RT process was understandable and correctly displayed from the children's point of view.

When observing the children playing, the main investigator noticed that the children spent more time on, for example, a dancing kitchen tap than on RT-related issues. This demonstrated that the children enjoyed the random nature of the interaction, and therefore obtained value from it; this is similar to an observation made by Howard-Jones and Demetriou [32], who found that players were more likely to prefer unpredicted rewards over ones they had foreseen. This gave us the idea to add previously unused scenes as hidden objects to create surprise and implement and test the addition. It may seem unconventional to depict an important procedure, like RT, in a playful way. However, as Clark [33] pointed out, resilient families and their ill children use imagination and humor as a way to cope with difficult situations. The amusement provided by the game did not directly contribute to education, but even if the learning aspects of the game became side impressions, the players still came in contact with them and learned from them. Mader et al [34] recommend that game designers focus on entertainment and fun when designing serious games to even

out the medical details. It needs to be noted that children and their parents preparing for RT are highly motivated to understand the procedure, or at least, not to be afraid of it. This creates a multi-faceted situation that is harder than usual to address with a serious game, whose function can normally be viewed as a process of convincing.

Learning Through Text

Using text as the means of learning in a serious game proved to have a roughly similar outcome as learning through game mechanics, as when using the latter, the player could misinterpret the message if they did not understand it correctly [35]. The use of metaphors to help children understand medical instructions in games has previously been tested, and the results show that children prefer different metaphors for practice [36]. In the current game, each medical phenomenon was explained by a single corresponding metaphor. The children said they understood the metaphors, but if the game is redesigned in the future, more metaphors might be added to further increase the children's ability to understand.

Limitations

Of the 14 families contacted, 9 chose to participate, including only 1 boy. While the data collected was rich, additional boys in the group could perhaps have provided additional information. One reason for the lack of boys was that at the time of inclusion, there were more girls than boys being treated with RT at the center. Two of the initial child participants left the study, one of whom said it was because they felt the game was more suited for younger children and one of whom did not give a reason. The data collection lasted for 8 months; thus, the procedure was time-consuming. However, since the children had different gaming backgrounds, they had different levels of background knowledge of computer and web games; this added richness to the way they comprehended the game under development that the investigators found more valuable than gender diversity. Further, these differences in gaming habits did not change when some of the children left the study. As suggested by Maheu-Cadotte et al [37], researchers should involve end users with varying gaming habits and use different methods to prompt their input when developing serious games. Using a qualitative method with several methods of data collection provided rich findings that strengthened the credibility of our conclusions and increased their transferability to similar areas under investigation [38]. Researchers use thematic analysis to develop themes that provide insight into the research question. It may be noted that the frequent concurrency of the ideas behind the themes in our data set has not been evaluated [39]; these data merely provide insight into individual experiences.

Future Research

Further evaluation of the game is warranted to determine whether it can increase knowledge of the procedure and thereby decrease anxiety. Thus, we are currently inviting children between 5 and 14 who are scheduled to undergo RT to evaluate the game.

Conclusion

Having children be part of the cocreation process through PAR resulted in several changes to the game. More importantly, the

methods used here made the children active participants; therefore, this method can be used by health care researchers to develop cocreated serious games together with children. When children participate in research, the study framework needs revision throughout the process, due to unforeseen

circumstances and in order to facilitate participation. It is necessary to inform the child participants that the process takes time, and that it can be revised to facilitate their participation as much as possible to avoid placing a burden on them.

Acknowledgments

We would like to acknowledge all the work performed by the game design team lead by Jakob Berglund Rogert at the Department of Game Design, Uppsala University. The authors warmly thank all the participating children and parents for the time and effort they put into the project. This project was funded by the Swedish Childhood Cancer Foundation.

Conflicts of Interest

None declared.

Multimedia Appendix 1

Semistructured interview guide.

[[DOCX File , 12 KB - humanfactors_v9i2e34476_app1.docx](#)]

Multimedia Appendix 2

Listed changes made to the game.

[[DOCX File , 16 KB - humanfactors_v9i2e34476_app2.docx](#)]

References

1. Hedström M, Haglund K, Skolin I, von Essen L. Distressing events for children and adolescents with cancer: child, parent, and nurse perceptions. *J Pediatr Oncol Nurs* 2003;20(3):120-132. [doi: [10.1053/jpon.2003.76](#)] [Medline: [12776260](#)]
2. Björk M, Wiebe T, Hallström I. Striving to survive: families' lived experiences when a child is diagnosed with cancer. *J Pediatr Oncol Nurs* 2005;22(5):265-275. [doi: [10.1177/1043454205279303](#)] [Medline: [16079362](#)]
3. Anderzén-Carlsson A, Kihlgren M, Svantesson M, Sorlie V. Parental handling of fear in children with cancer; caring in the best interests of the child. *J Pediatr Nurs* 2010 Oct;25(5):317-326. [doi: [10.1016/j.pedn.2008.10.004](#)] [Medline: [20816553](#)]
4. Carleton RN. Fear of the unknown: One fear to rule them all? *J Anxiety Disord* 2016 Jun;41:5-21 [[FREE Full text](#)] [doi: [10.1016/j.janxdis.2016.03.011](#)] [Medline: [27067453](#)]
5. Woodman H. Put yourself in their shoes ... the vulnerability of children and their families when attending for radiotherapy treatment: The role of the specialist paediatric radiotherapy radiographer. *Radiography* 2013 Nov;19(4):311-314. [doi: [10.1016/j.radi.2013.08.006](#)]
6. Klosky J, Tyc VL, Tong X, Srivastava DK, Kronenberg M, de Armendi AJ, et al. Predicting pediatric distress during radiation therapy procedures: the role of medical, psychosocial, and demographic factors. *Pediatrics* 2007 May;119(5):e1159-e1166. [doi: [10.1542/peds.2005-1514](#)] [Medline: [17452492](#)]
7. Gårdling J, Törnqvist E, Månsson ME, Hallström IK. Age-appropriate preparations for children with cancer undergoing radiotherapy: A feasibility study. *J Child Health Care* 2017 Dec;21(4):370-380. [doi: [10.1177/1367493517727070](#)] [Medline: [29110518](#)]
8. Filin A, Treisman S, Peles Bortz A. Radiation therapy preparation by a multidisciplinary team for childhood cancer patients aged 3 1/2 to 6 years. *J Pediatr Oncol Nurs* 2009;26(2):81-85. [doi: [10.1177/1043454208328766](#)] [Medline: [19202117](#)]
9. Clerici C, Ferrari A, Pecori E, Diletto B, Alessandro O, Pagani Bagliacca E, et al. Age-appropriate multidisciplinary approach to young children with cancer undergoing radiotherapy: The SIESTA procedure. *Pediatr Blood Cancer* 2021 Jan;68(1):e28650. [doi: [10.1002/pbc.28650](#)] [Medline: [32779871](#)]
10. Engvall G, Lindh V, Mullaney T, Nyholm T, Lindh J, Ångström-Brännström C. Children's experiences and responses towards an intervention for psychological preparation for radiotherapy. *Radiat Oncol* 2018 Jan 22;13(1):9 [[FREE Full text](#)] [doi: [10.1186/s13014-017-0942-5](#)] [Medline: [29357940](#)]
11. Plass JL, Homer BD, Kinzer CK. Foundations of Game-Based Learning. *Educ Psychol* 2016 Feb 06;50(4):258-283. [doi: [10.1080/00461520.2015.1122533](#)]
12. Rittel HWJ, Webber MM. Dilemmas in a general theory of planning. *Policy Sci* 1973 Jun;4(2):155-169. [doi: [10.1007/bf01405730](#)]
13. Zavcer G, Mayr S, Petta P. Design Pattern Canvas: An Introduction to Unified Serious Game Design Patterns. *Interdiscip. Descr Complex Syst* 2014;12(4):280-292. [doi: [10.7906/indecs.12.4.2](#)]
14. Adams E. *Fundamentals of game design*, 3rd ed. Berkeley, CA: New Riders; 2014.

15. Wattanasoontorn V, Boada I, García R, Sbert M. Serious games for health. *Entertain Comput* 2013 Dec;4(4):231-247. [doi: [10.1016/j.entcom.2013.09.002](https://doi.org/10.1016/j.entcom.2013.09.002)]
16. Liszio S, Basu O, Masuch M. A Universe Inside the MRI Scanner: An In-Bore Virtual Reality Game for Children to Reduce Anxiety and Stress. In: CHI PLAY '20: Proceedings of the Annual Symposium on Computer-Human Interaction in Play. 2020 Presented at: CHI PLAY: Computer-Human Interaction in Play; Nov 2-4, 2020; Virtual Event, Canada. [doi: [10.1145/3410404.3414263](https://doi.org/10.1145/3410404.3414263)]
17. Elsbernd A, Hjerding M, Visler C, Hjalgrim LL, Niemann CU, Boisen KA, et al. Using Cocreation in the Process of Designing a Smartphone App for Adolescents and Young Adults With Cancer: Prototype Development Study. *JMIR Form Res* 2018 Nov 01;2(2):e23 [FREE Full text] [doi: [10.2196/formative.9842](https://doi.org/10.2196/formative.9842)] [Medline: [30684439](https://pubmed.ncbi.nlm.nih.gov/30684439/)]
18. Verschuere S, Buffel C, Vander Stichele G. Developing Theory-Driven, Evidence-Based Serious Games for Health: Framework Based on Research Community Insights. *JMIR Serious Games* 2019 May 02;7(2):e11565 [FREE Full text] [doi: [10.2196/11565](https://doi.org/10.2196/11565)] [Medline: [31045496](https://pubmed.ncbi.nlm.nih.gov/31045496/)]
19. White G, Suchowierska M, Campbell M. Developing and systematically implementing participatory action research. *Arch Phys Med Rehabil* 2004 Apr;85(4 Suppl 2):S3-12. [doi: [10.1016/j.apmr.2003.08.109](https://doi.org/10.1016/j.apmr.2003.08.109)] [Medline: [15083417](https://pubmed.ncbi.nlm.nih.gov/15083417/)]
20. Langhout R, Thomas E. Imagining participatory action research in collaboration with children: an introduction. *Am J Community Psychol* 2010 Sep;46(1-2):60-66 [FREE Full text] [doi: [10.1007/s10464-010-9321-1](https://doi.org/10.1007/s10464-010-9321-1)] [Medline: [20571860](https://pubmed.ncbi.nlm.nih.gov/20571860/)]
21. Ingadottir B, Laitonen E, Stefansdottir A, Sigurdardottir AO, Brynjolfsdottir B, Parisod H, et al. Developing a Health Game to Prepare Preschool Children for Anesthesia: Formative Study Using a Child-Centered Approach. *JMIR Serious Games* 2022 Jan 20;10(1):e31471 [FREE Full text] [doi: [10.2196/31471](https://doi.org/10.2196/31471)] [Medline: [35049507](https://pubmed.ncbi.nlm.nih.gov/35049507/)]
22. Zimmerman J, Forlizzi J, Evenson S. Research through design as a method for interaction design research in HCI. In: Proceedings of the SIGCHI conference on Human factors in computing systems. 2007 Presented at: CHI07: CHI Conference on Human Factors in Computing Systems; Apr 28-May 3, 2007; San Jose, CA. [doi: [10.1145/1240624.1240704](https://doi.org/10.1145/1240624.1240704)]
23. Braun V, Clarke V. Using thematic analysis in psychology. *Qual Res Psychol* 2006 Jan;3(2):77-101 [FREE Full text] [doi: [10.1191/1478088706qp0630a](https://doi.org/10.1191/1478088706qp0630a)]
24. Martey RM, Stromer-Galley J. The Digital Dollhouse: Context and Social Norms in The Sims Online. *Games Cult* 2016 Apr 22;2(4):314-334. [doi: [10.1177/1555412007309583](https://doi.org/10.1177/1555412007309583)]
25. Hunicke R, LeBlanc M, Zubek R. MDA: A formal approach to game design and game research. Association for the Advancement of Artificial Intelligence. URL: <https://www.aaai.org/Papers/Workshops/2004/WS-04-04/WS04-04-001.pdf> [accessed 2022-05-24]
26. Nygren JM, Lindberg S, Wärnestål P, Svedberg P. Involving Children With Cancer in Health Promotive Research: A Case Study Describing Why, What, and How. *JMIR Res Protoc* 2017 Feb 07;6(2):e19 [FREE Full text] [doi: [10.2196/resprot.7094](https://doi.org/10.2196/resprot.7094)] [Medline: [28174150](https://pubmed.ncbi.nlm.nih.gov/28174150/)]
27. Bracken-Roche D, Bell E, Macdonald ME, Racine E. The concept of 'vulnerability' in research ethics: an in-depth analysis of policies and guidelines. *Health Res Policy Syst* 2017 Feb 07;15(1):8 [FREE Full text] [doi: [10.1186/s12961-016-0164-6](https://doi.org/10.1186/s12961-016-0164-6)] [Medline: [28173859](https://pubmed.ncbi.nlm.nih.gov/28173859/)]
28. Shamrova DP, Cummings CE. Participatory action research (PAR) with children and youth: An integrative review of methodology and PAR outcomes for participants, organizations, and communities. *Child Youth Serv Rev* 2017 Oct;81:400-412. [doi: [10.1016/j.chidyouth.2017.08.022](https://doi.org/10.1016/j.chidyouth.2017.08.022)]
29. Ruland CM, Starren J, Vatne TM. Participatory design with children in the development of a support system for patient-centered care in pediatric oncology. *J Biomed Inform* 2008 Aug;41(4):624-635 [FREE Full text] [doi: [10.1016/j.jbi.2007.10.004](https://doi.org/10.1016/j.jbi.2007.10.004)] [Medline: [18082468](https://pubmed.ncbi.nlm.nih.gov/18082468/)]
30. Shier H. Pathways to participation: openings, opportunities and obligations. *Child Soc* 2001 Apr;15(2):107-117. [doi: [10.1002/chi.617](https://doi.org/10.1002/chi.617)]
31. Thomas F, Johnston O. The principles of animation. In: *Disney animation: the illusion of life*. New York, NY: Abbeville Press; 1981:47-70.
32. Howard-Jones PA, Demetriou S. Uncertainty and engagement with learning games. *Instr Sci* 2008 Sep 11;37(6):519-536. [doi: [10.1007/s11251-008-9073-6](https://doi.org/10.1007/s11251-008-9073-6)]
33. Clark C. Imaginal coping: Resilience through a play of tropes. In: Clark C, Ferrari M, editors. *Child and Adolescent Resilience Within Medical Contexts*. Cham, Switzerland: Springer; 2016:175-192.
34. Mader S, Natkin S, Levieux G. How to Analyse Therapeutic Games: The Player / Game / Therapy Model. In: *Entertainment Computing - ICEC 2012. Lecture Notes in Computer Science*. Berlin, Germany: Springer; 2012 Presented at: International Conference on Entertainment Computing; Sep 26-29, 2012; Bremen, Germany. [doi: [10.1007/978-3-642-33542-6_17](https://doi.org/10.1007/978-3-642-33542-6_17)]
35. Molnar A, Kostkova P. On effective integration of educational content in serious games: Text vs. game mechanics. 2013 Presented at: IEEE 13th International Conference on Advanced Learning Technologies; Jul 15-18, 2013; Beijing, China. [doi: [10.1109/icalt.2013.94](https://doi.org/10.1109/icalt.2013.94)]
36. Delden R, Plass-Oude Bos D, de With AJAV, Vogel K, Klaassen R, Zwart N, et al. SpiroPlay, a Suite of Breathing Games for Spirometry by Kids & Experts. In: CHI PLAY '20: Proceedings of the Annual Symposium on Computer-Human Interaction in Play. 2020 Presented at: Annual Symposium on Computer-Human Interaction in Play; Nov 2-4, 2020; Virtual Event, Canada p. 400. [doi: [10.1145/3410404.3414223](https://doi.org/10.1145/3410404.3414223)]

37. Maheu-Cadotte M, Dubé V, Cossette S, Lapierre A, Fontaine G, Deschênes MF, et al. Involvement of End Users in the Development of Serious Games for Health Care Professions Education: Systematic Descriptive Review. *JMIR Serious Games* 2021 Aug 19;9(3):e28650 [FREE Full text] [doi: [10.2196/28650](https://doi.org/10.2196/28650)] [Medline: [34129514](https://pubmed.ncbi.nlm.nih.gov/34129514/)]
38. Shenton AK. Strategies for ensuring trustworthiness in qualitative research projects. *Educ Inf* 2004 Jul 19;22(2):63-75. [doi: [10.3233/efi-2004-22201](https://doi.org/10.3233/efi-2004-22201)]
39. Kiger ME, Varpio L. Thematic analysis of qualitative data: AMEE Guide No. 131. *Med Teach* 2020 Aug;42(8):846-854. [doi: [10.1080/0142159X.2020.1755030](https://doi.org/10.1080/0142159X.2020.1755030)] [Medline: [32356468](https://pubmed.ncbi.nlm.nih.gov/32356468/)]

Abbreviations

PAR: participatory action research

RT: radiotherapy

Edited by A Kushniruk; submitted 27.10.21; peer-reviewed by R Vilela, T Silva; comments to author 03.02.22; revised version received 22.03.22; accepted 19.04.22; published 31.05.22.

Please cite as:

Cederved C, Back J, Ångström-Brännström C, Ljungman G, Engvall G

Co-creation of a Serious Game About Radiotherapy: Participatory Action Research Study With Children Treated for Cancer

JMIR Hum Factors 2022;9(2):e34476

URL: <https://humanfactors.jmir.org/2022/2/e34476/>

doi: [10.2196/34476](https://doi.org/10.2196/34476)

PMID: [35639467](https://pubmed.ncbi.nlm.nih.gov/35639467/)

©Catarina Cederved, Jon Back, Charlotte Ångström-Brännström, Gustaf Ljungman, Gunn Engvall. Originally published in *JMIR Human Factors* (<https://humanfactors.jmir.org>), 31.05.2022. This is an open-access article distributed under the terms of the Creative Commons Attribution License (<https://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in *JMIR Human Factors*, is properly cited. The complete bibliographic information, a link to the original publication on <https://humanfactors.jmir.org>, as well as this copyright and license information must be included.

Original Paper

Development of a Digital Tool for People With a Long-Term Condition Using Stroke as a Case Example: Participatory Design Approach

Emma K Kjörk^{1,2}, MSc; Katharina S Sunnerhagen^{1,2}, PhD; Åsa Lundgren-Nilsson^{1,2}, PhD; Anders K Andersson³; Gunnel Carlsson¹, PhD

¹Department of Clinical Neuroscience, Institute of Neuroscience and Physiology, Sahlgrenska Academy, University of Gothenburg, Gothenburg, Sweden

²Sahlgrenska University Hospital, Gothenburg, Sweden

³Stroke Association, Gothenburg, Sweden

Corresponding Author:

Emma K Kjörk, MSc

Department of Clinical Neuroscience, Institute of Neuroscience and Physiology, Sahlgrenska Academy

University of Gothenburg

Per Dubbsgatan 14, 3 floor

Gothenburg, 413 45

Sweden

Phone: 46 735124309

Email: emma.kjork@gu.se

Abstract

Background: In patient care, demand is growing for digital health tools to enable remote services and enhance patient involvement. People with chronic conditions often have multiple health problems, and long-term follow-up is recommended to meet their needs and enable access to appropriate support. A digital tool for previsit preparation could enhance time efficiency and guide the conversation during the visit toward the patient's priorities.

Objective: This study aims to develop a digital previsit tool and explore potential end user's perceptions, using a participatory approach with stroke as a case example.

Methods: The digital tool was developed and prototyped according to service design principles, informed by qualitative participant data and feedback from an expert panel. All features were processed in workshops with a team that included a patient partner. The resulting tool presented questions about health problems and health information. Study participants were people with stroke recruited from an outpatient clinic and patient organizations in Sweden. Development and data collection were conducted in parallel. For conceptualization, the initial prototype was based on the Post-Stroke Checklist and research. Needs and relevance were explored in focus groups, and we used a web survey and individual interviews to explore perceived utility, ease of use, and acceptance. Data were thematically analyzed following the Framework Method.

Results: The development process included 22 participants (9 women) with a median age of 59 (range 42-83) years and a median of 51 (range 4-228) months since stroke. Participants were satisfied or very satisfied with using the tool and recommended its use in clinical practice. Three main themes were constructed based on focus group data (n=12) and interviews (n=10). First, valuable accessible information illuminated the need for information to confirm experiences, facilitate responses, and invite engagement in their care. Amendments to the information in turn reconfigured their expectations. Second, utility and complexity in answering confirmed that the questions were relevant and comprehensible. Some participants perceived the answer options as limiting and suggested additional space for free text. Third, capturing needs and value of the tool highlighted the tool's potential to identify health problems and the importance of encouraging further dialog. The resulting digital tool, Strokehälsa [Strokehealth] version 1.0, is now incorporated into a national health platform.

Conclusions: The participatory approach to tool development yielded a previsit digital tool that the study group perceived as useful. The holistic development process used here, which integrated health information, validated questions, and digital functionality, offers an example that could be applicable in the context of other long-term conditions. Beyond its potential to identify care needs, the tool offers information that confirms experiences and supports answering the questions in the tool. The tool is freely shared for adaptation in different contexts.

Trial Registration: researchweb 236341; <https://www.researchweb.org/is/vgr/project/236341>

(*JMIR Hum Factors* 2022;9(2):e35478) doi:[10.2196/35478](https://doi.org/10.2196/35478)

KEYWORDS

eHealth; digital tool; Strokehälsa; follow-up; chronic care; stroke; Post-Stroke Checklist; health literacy; co-design; shared decision making

Introduction

The desire to encourage patient involvement [1] and the growing acceptance of digitized health care have contributed to a rising interest in digital health tools [2,3]. For people living with long-term conditions [4] and chronic disability [5], recurrent interventions and health care support are crucial. Stroke, which serves as the context of this study, is a common cause of disability, with more than 101 million cases worldwide in 2019 [6], and is associated with motor impairments and cognitive, emotional, and communication difficulties [5]. Organized systems of care, including follow-up and self-management, are beneficial for people with long-term conditions [4,7], particularly because they can have more difficulty in actively engaging with health services [8-10]. A redesign of health care services to ensure a prepared patient and a proactive health care team is crucial [4]. Digital health solutions may offer tools that can facilitate improved follow-up.

Digitizing has already been shown to speed up the redesign of health care [3] and potentially foster access to health care services [7,11]. In Sweden, the digital platform *Healthcare Guide 1177* [12], accessible to registered individuals, is widely used. The platform includes medical information and health tools such as previsit forms. However, regardless of the mode of service delivery, patients need person-centered support [1] and accessible health information [10]. Thus, digital tool design should take into consideration patient need for information [13] and the best ways to promote active engagement of patients in long-term care with their health professionals [2].

The use of previsit tools can make people feel more knowledgeable, better informed, and clearer about their values [14]. Moreover, digital health tools can enhance dialog with health professionals [15] and empower patients to become active partners [2]. A recent randomized clinical trial showed that a previsit digital tool for collecting contextual data from patients had a positive impact on patient-provider communication [16]. However, previsit digital tools often focus on collecting data but do not include health information for the patient [12,16,17], even though information is key to eHealth literacy [10], comprising a patient's ability to understand, access, and use eHealth technologies. Patients need to be involved in self-management and interactions with health professionals [9,18], and digital solutions must be designed to promote eHealth literacy [10], counteract inequalities [18], and enhance shared decision making [19]. A thorough design process is indispensable to achieving this aim.

Service design is a human-centered approach that focuses on understanding the patient experiences to achieve a holistic view of solutions to complex problems [19,20]. Different methods

can be applied, such as a “persona” that represents a member of a future user group, or a set of prototypes that offer alternative solutions for a digital tool [20,21]. Qualitative research often can be used to explore the needs of patients and health care providers before pilot versions of such tools are tested [20] or incorporated into secure health platforms for use. A combination of service design and co-design approaches is beneficial for understanding users' needs in terms of technologies or processes [22]. In participatory co-design approaches, stakeholders—such as researchers, patients, and health care staff—work together throughout the design process [23]. The Technological Acceptance Model (TAM) [24] illustrates factors influencing adoption of technology and how perceived usefulness and ease of use affect acceptability.

People with long-term conditions need digital tools designed to cover a range of health problems and related information. Although previsit digital tools have been designed for people with various conditions [2,16,17,25], tools related to organized follow-up after stroke are scarce [11]. Furthermore, when digital elements are used, they are commonly part of a comprehensive and complex approach to poststroke follow-up [26,27] and lack a thorough description of the development process including user experiences [26]. Additionally, tools commonly request patient-reported data [16,17] without a combined solution in which patients in turn receive tailored health information. To our knowledge, no user-friendly previsit digital tool is yet available that includes well-validated self-report and health information to prepare people with stroke for a follow-up visit. Our aim was to develop a digital previsit tool and explore potential end user's perceptions prior to testing it in a clinical setting, using a participatory approach with stroke as a case example.

Methods

Overview

A participatory [23] and pragmatic approach including mixed methods [28] was used to design a digital tool that meets users' needs. In a participatory co-design approach, end users are viewed as experts on their experiences, and they can be engaged at different levels, with some becoming partners in the research team [23,29]. To ensure patient involvement, a patient with stroke who was engaged in a support association and had a background in information technology projects became a patient partner and coauthor. Initially, this patient partner (AKA) provided valuable advice regarding recruitment, participant involvement, and how to introduce the prototype. He reviewed the tool content and was involved in workshops with members of the research team (EKK, GC, and KSS). He also reviewed the summaries of the preliminary themes.

Participants and Recruitment

Participants were included in 2 phases between December 2017 and October 2020. Staff identified eligible individuals in 3 settings: an outpatient unit, Stroke Forum (a center for support and advice after initial care), and a support association. Potential participants were briefly informed about the study, and those who agreed to participate were contacted by a researcher (EKK) via phone to provide detailed information. After purposive sampling with an attempt to achieve heterogeneity in terms of age, sex, communication, and mobility, participants were scheduled for an interview. They also were sent study information by email, including web links to a pilot version of the tool and to a web survey for the amendment phase. The inclusion criterion was having had a stroke. The exclusion criterion was severe communication or cognitive difficulties that made participation impossible, even in a small group discussion or together with next of kin. The sample size was guided by the concept of information power to enhance the richness of data according to the aim [30]. All participants gave their informed consent before data collection.

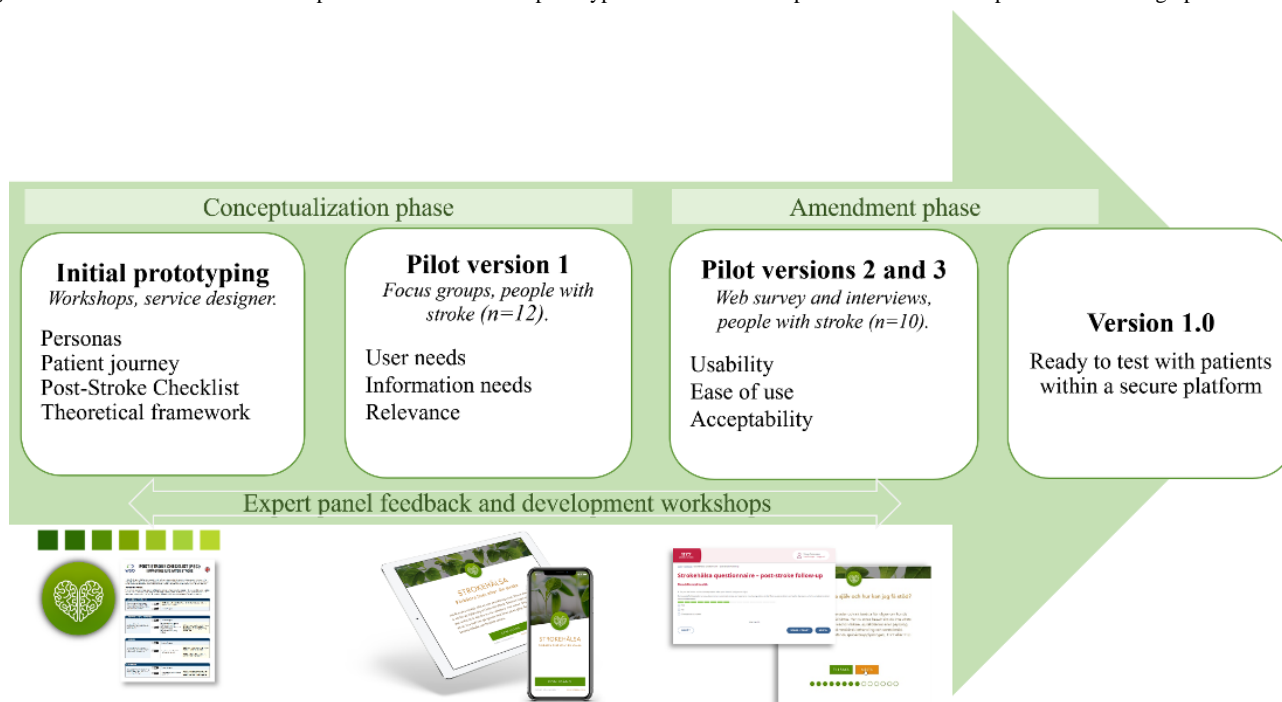
Members of the expert panel were contacted by the first author throughout the process during 2017-2020. Potential members

were purposively recruited to represent different services, including members of the stroke association and health care professionals. Expert panel members received the link to the second pilot version and a separate MS Word document with the text included in the tool. Written feedback on text revisions was collected via email, unless verbal input was preferred. Members of relevant professions then were approached for specific feedback when appropriate. The expert panel (n=11, 3 males) had a median age of 55 years (range 42-70 years) and represented the following competencies: nurse (n=1), occupational therapist (n=3), physician (n=1), physiotherapist (n=1), neuropsychologist (n=1), speech therapist (n=2), service designer (n=1), and patient partner (n=1). Professionals had a median of 20 years (range 10-40 years) of stroke experience with the following education levels: doctoral degree (PhD; n=2), PhD student (n=3), and master's degree (n=4).

Development and Data Collection

The co-design process integrated the development of the tool and the data collection, including user experiences. The process was performed in 2 phases: the conceptualization phase and the amendment phase, including a variety of methods involving different stakeholders (Figure 1).

Figure 1. The different sources for input and visualization of prototypes used in the conceptual and amendment phases of the design process.



Development

The starting point for the conceptual phase comprised previous research regarding follow-up and suggestions for a patient version of the Post-Stroke Checklist (PSC) [31]. It was specified in advance that a digital previsit tool should be developed based on the PSC. Existing research related to stroke and person-centered care informed this development to ensure an evidence-based tool. The initial prototyping workshops were conducted by a service designer and the first author (EKK). The PSC [31,32], which constituted the basis for mapping the “patient journey” [20], is an easy tool for identifying common

health problems and facilitating further actions, such as referrals to health services or patient organizations. It comprises 11 questions within the following areas: secondary prevention, activities of daily living, mobility, spasticity, pain, incontinence, communication, mood, cognition, life after stroke, and relationship with family. One example of a question is, “Since your stroke or last assessment, are you finding it more difficult to communicate with others?” The PSC is available from the World Stroke Organization and free for anyone to use. To foster an understanding of the potential user group, “personas” [20] were developed to represent users of different sexes, ages, personalities, life situations, values, and interests.

The content of and amendments to the tool were discussed during formal workshops and in dialogs with the research team, including the patient partner (EKK, GC, KSS, and AKA). These workshops guided the design of new pilot versions, and decisions were taken in consensus. Decisions were based on the data collection, expert panel feedback, and relevant evidence, and addressed, for example, clarification of the language by the addition or removal of text and answer options. All relevant data were combined into a single document before being systematically discussed in the workshops.

Data Collection: User Needs and Experiences

During the conceptualization phase from December 2017 to March 2019, focus group discussions [33] were conducted within each location to explore user needs and the perceived relevance of an early prototype, known as pilot 1. Participants tested the pilot in their home environment before the focus group took place. In all groups, an interview guide [33] with open-ended questions was used (Multimedia Appendix 1), and the moderator gave a summary at the end, leaving open the possibility of correcting potential misunderstandings. Each focus group was conducted face-to-face, lasted approximately 60-90 minutes, and was audio recorded and transcribed verbatim. The focus group data and field notes informed further amendments and prototyping.

In the amendment phase from September to October 2020, pilot versions 2 and 3 were tested before data collection. Participants completed a web survey, followed by an individual interview, to explore perceived usefulness, ease of use, and acceptability, as inspired by the TAM [24]. The survey included previous web habits, demographic data, and satisfaction ratings, as follows: very satisfied, satisfied, dissatisfied, very dissatisfied, and don't know. Along with the interview, self-reported characteristics were collected, including stroke type, time since stroke, and level of dependence in activities of daily living. The interviews were conducted by the first author (EKK) via phone, who followed an interview guide (Multimedia Appendix 1). During the interview, participants did "think aloud" [34] as they were using the tool. Support from next of kin was allowed during data collection.

Analysis

Substantive cross-sectional data analysis was performed, in line with the Framework Method [35]. In accordance with this pragmatic approach, the analysis combined data from focus groups and individual interviews and involved 5 steps, performed mainly by the first author (EKK) in cooperation with

the last author (GC). In the *first step*, all transcribed interviews were read to achieve familiarization and get an overview of the content. In the *second step*, an initial framework was constructed based on the different parts of the tool (ie, information and answering), which was then revised after the first interviews. In the *third step*, the transcribed data were indexed according to "codes" and sorted based on the initial framework. NVivo 11 software was used for data management. In the *fourth step*, the data extracts were reviewed together to ensure that similar content was sorted together and to determine whether the theme titles should be adapted. In the *fifth step*, data were summarized and displayed in a matrix in an MS Word file. Each subtheme was summarized based on the codes and raw data. The individual interviews, combined, were handled as one case and each focus group as separate cases. Data from each case was summarized separately before all cases were merged. The systematic data management using NVivo enabled easy access back to the initial subthemes and interview transcripts. Finally, the patient partner and coauthors read the summaries and were involved in refining the themes. Throughout the process, memos were written to summarize reflections, alternative interpretations, and potential amendments. Data collected from the web survey and self-reported characteristics were analyzed using descriptive statistics and are presented as numbers or medians and ranges. Analyses were performed using SPSS version 24 (IBM, Inc.).

Ethical Approval

The study was approved by the Swedish Ethical Review Authority (no. 556-17 and 2020-03324).

Results

Design

The participatory design process grounded in user experiences resulted in a digital previsit tool. The following description of participants, the process for development, and user experiences provide insights into the rationale for amendments that were made.

Participants

The study included a total of 22 individuals with stroke (9 women), with a median age of 59 (range 42-83) years, and a median of 51 (range 4-228) months since stroke onset. Together, the participants represented a wide range of individual characteristics (Table 1).

Table 1. Characteristics of participants included in the study.

Participants	Focus group 1 (n=3)	Focus group 2 (n=4)	Focus group 3 (n=5)	Interview, survey (n=10) ^a
Age, median (range)	67 (64-83)	65 (43-73)	55 (47-70)	54 (42-74)
Male sex, n	1	2	4	6
Education (highest degree), n				
Mandatory	0	0	1	0
High school	2	1	4	7
University	1	3	0	3
Source of income, n				
Work	0	0	1	6
Sick leave	1	2	2	2
Retirement	2	2	2	2
Months since stroke, median (range)	192 (168-228)	126 (72-156)	14 (4-24)	42 (13-144)
Stroke characteristics (self-report), n				
Ischemic stroke	3	3	4	7
Location				
Right	0	2	3	5
Left	2	1	2	3
Posterior	1	1	0	2
Communication difficulties (aphasia), n	2	1	2	5
Activities of daily living independency, n	2	4	4	8
Internet use daily, n	1	4	5	8

^aOne participant (male) answered the web survey but did not participate in an interview afterward.

Development

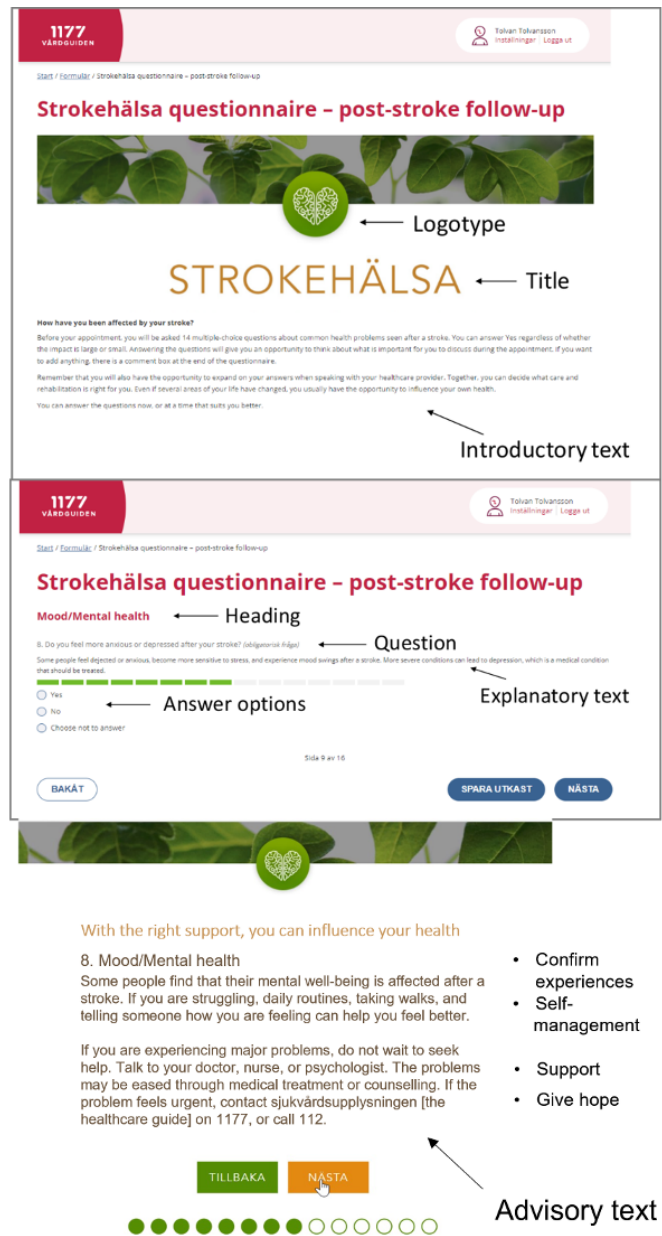
The development process included 3 pilot versions of the tool (Figure 1) and stepwise alterations (Multimedia Appendix 2) before version 1.0 was completed (for visual presentation, see Multimedia Appendix 3).

In the conceptualization phase, the initial prototypes (webpages) were developed in collaboration with the first author and a service designer. During this stage, the focus was on user's needs rather than fitting into a specific platform. The first digital pilot, version 1, included the following components: a logotype and title Strokehälsa [Strokehealth], introductory information, questions about health problems to be answered with yes or no, explanatory text (linked to "read more") placed in direct connection with each question, and summary of results. The questions were in accordance with the PSC when applicable. The name Strokehälsa was chosen based on the aim of promoting health and improved life after stroke. The layout was intended to be clean and to avoid overwhelming text while remaining open to the possible addition of more information in the future. The information was layered using "read more" texts, with the aim of adapting the information level to each individual. The explanatory (read more) texts were inspired by existing

patient information, such as pamphlets and booklets, as a starting point for gathering opinions.

In the amendment phase, alterations were performed based on preliminary findings from the focus groups and the theoretical framework. The integration between the central components, validated questions, health information, and technical aspects was essential to improve usability. Thus, beyond revising the questions, adding real-life examples in the explanatory text enhanced usability. Important changes in pilot 2 were the inclusion of advisory texts, with brief information about support options and self-management, and a free textbox offering the opportunity to describe "other challenges." In pilot 3, a general question about rehabilitation and a place for free-text comments were added. This pilot was incorporated into the national platform. However, the platform has some limitations regarding layout options, such as no hidden read more text option and predefined typography and colors. Figure 2 shows a screenshot of the patient view of the tool. The functionality of the platform includes sending the patient an email or SMS text message notification with an invitation to use the tool and to answer the questions before a care visit. Responsible health professionals can send version 1.0 of the tool to patients and view the summary of results (using a staff log-in at the secure platform 1177) before a care visit.

Figure 2. Screenshot with an overview of the core functions in the tool.



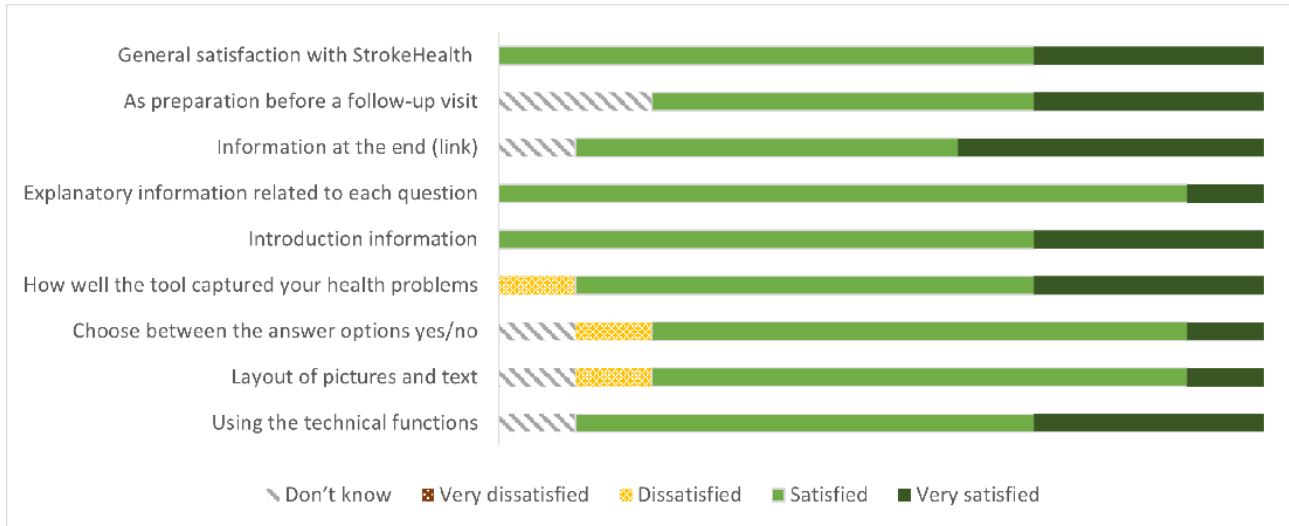
User Needs and Experiences

Results regarding user experiences were based on the satisfaction survey and qualitative interviews.

Satisfaction Survey

Satisfaction with the tool was high (Figure 3). Most participants were satisfied, and all participants in the amendment phase would recommend use of the tool in clinical practice. Participants used different devices, with the majority using their mobile phone (n=16), followed by tablet (n=3), computer (n=2), and more than 1 device (n=1).

Figure 3. Satisfaction with using the tool (n=10).



Themes Created Based on Focus Groups and Individual Interviews

Overview

Analyses of focus group discussions and individual interviews were combined to clarify the meaning of the users’ experiences

with the tool. Experiences were summarized in the overarching theme (*A multifaceted digital solution—essential to empower patients before a care visit*), main themes, subthemes (Figure 4), and quotes (Table 2).

Figure 4. Themes and subthemes based on user experiences described in focus groups and individual interviews. a: Only data from focus groups; b: only data from individual interviews.

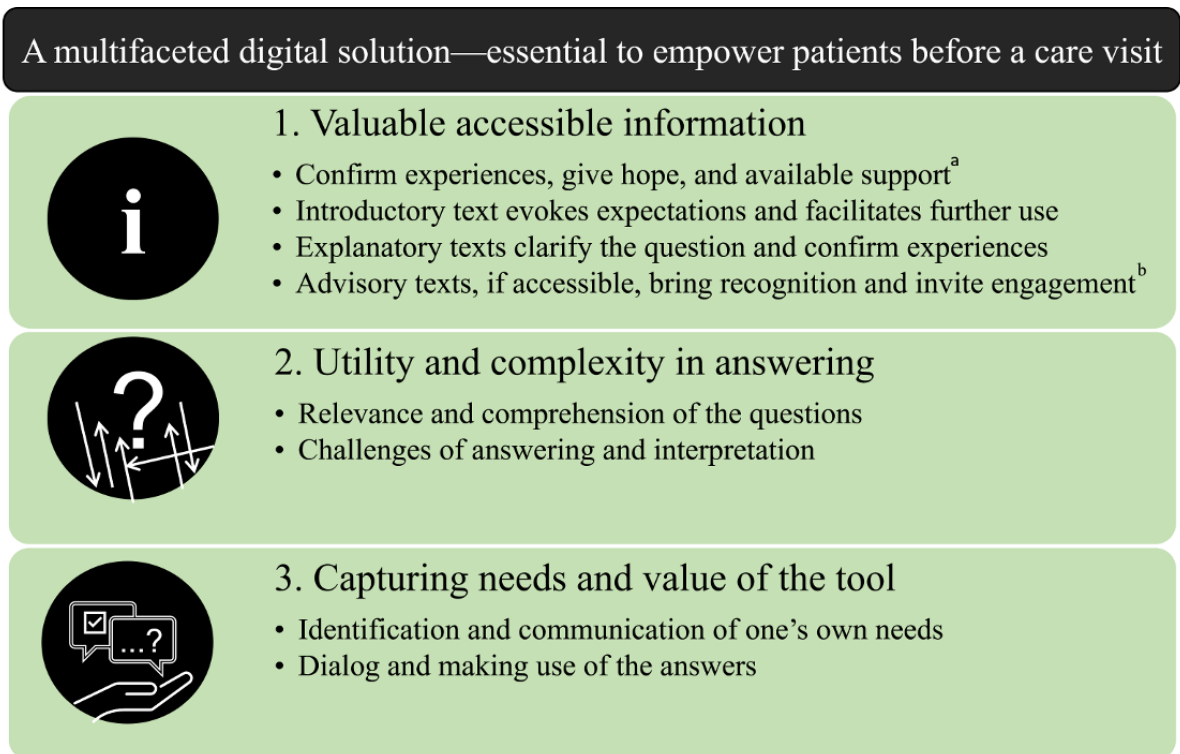


Table 2. Quotes representing the themes in the qualitative analyses based on focus groups and individual interviews.

Themes and subthemes	Quotes
Theme 1: Valuable accessible information	
Confirm experiences, give hope, and available support	<p>P1^a: <i>It's like that</i></p> <p>Moderator: <i>That you're not alone - is that what you want to communicate/say?</i></p> <p>P1: <i>Very important</i></p> <p>P2: <i>Yep</i></p> <p>Moderator: <i>(it's) not just you?</i></p> <p>P3: <i>It's good to know you aren't alone</i></p> <p>[Focus group 1, people with communication difficulties]</p>
Introductory text evokes expectations and facilitates further use	<p><i>Yes, I expect that I can find out what I can do for myself to improve my health, as written now (introductory text), it provides an expectation but that wasn't fulfilled. (Reads advisory texts). (...) Yeah yeah, but I get it, now that I've been able to read your link, why what is written there, is written there. [P4, woman, 50 years old]</i></p>
Explanatory texts clarify the question and confirm experiences	<p><i>I think that the "read more" gave me enough information to be able to maybe change my answer to yes rather than no. [P5, Focus group 2]</i></p> <p><i>There was a little bit more (information) there and I think that it was easier to answer, like "yeah, that fits really well with the answer I want to give", so I didn't have to doubt (my answer). [P6, man, 42 years old]</i></p>
Advisory texts, if accessible, bring recognition and invite engagement	<p><i>I don't think I got it, no I didn't get it clearly actually (...) I was like, kinda done with the questionnaire and so I was done. [P7, woman, 53 years old]</i></p> <p><i>Yes, I see it. Much easier to read (divided in sections and Strokehälsa layout) yes, it's pretty fantastic really but, it's actually the case, that I found it to be, that it was more interesting to read it that way. Because then you were a bit curious about what was on the next page, otherwise you see all the titles at once and then it's, it looks like a lot, that this didn't look like too much. [P8, man, 66 years old]</i></p>
Theme 2: Utility and complexity in answering	
Relevance and comprehension of the questions	<p><i>I think that the questions were good, easy to answer, I didn't think it was a problem for me to answer them. So, I didn't see any obvious gaps, like "oh this, I didn't get this or didn't understand that."</i></p> <p>[P6, man, 42 years old].</p> <p>P5: <i>Yeah, I never got into those questions if I answered no (Activities of daily living).</i></p> <p>Moderator: <i>Do you think you miss out on people who don't end up answering the follow up questions if they main questions are too broad?</i></p> <p>P5: <i>That can happen.</i></p> <p>P9: <i>Yes, I think so (too) because I just saw that if you click on yes there are more questions that come up.</i></p> <p>P10: <i>And yeah, it's easy to press no.</i></p> <p>[Focus group 2]</p>
Challenge of answering and interpretation	<p><i>No, I think that some of the questions could have been so that you only had one choice, I told you a bunch of times that they're on the edge, that if you could squeeze in a third option, so it's yes or no, there's something in between, you heard I was in a grey zone several times (between yes and no), you can maybe put it like that.</i></p> <p>[P11, man, 59 years old]</p> <p><i>It would be if you could add some kind of comment somewhere. Because if none of the questions are appropriate you could just write something yourself. But that's usually the problem, that you don't write anything yourself, just answer yes/no, but the opportunity to write something would be good.</i></p> <p>[P12, man, 56 years old]</p>
Theme 3: Capturing needs and value of the tool	

Themes and subthemes	Quotes
Identification and communication of one's own needs	<p>P13: <i>It could be a support by prompting certain questions that you maybe hadn't thought of</i></p> <p>P14: <i>Or it prompts the questions that you'd thought of but had kind of just swept under the carpet. (...) Getting these questions and thoughts aired so that you can get your thing moving</i></p> <p>P15: <i>You can get answers to questions you maybe didn't understand or didn't get answers to.</i> [Focus group 3]</p> <p><i>If it's there (in the tool), maybe you'll be brave enough to bring it up (sex), with your doctor. Otherwise, it might be a bit too sensitive to mention it.</i> [P8, man, 66 years old]</p>
Dialog and making use of the answers	<p><i>No, it's enough that you just answer with a yes, if that yes gets some attention at a doctor's appointment.</i> [P7, Woman, 53 years old]</p> <p><i>I'm not saying that I speak for everyone, but I think that lot of people want to have the possibility, at least, to tell someone how you feel, that meeting people in between is preferable.</i> [P16, man, 51 years old]</p>

^aP: participant.

Valuable Accessible Information

Overview

This theme comprises the perceived value of information to confirm experiences after a stroke episode and to facilitate further use of the tool.

Confirm Experiences, Give Hope, and Available Support

Participants emphasized their need for reliable and targeted information. Participants strongly agreed that they wanted information to confirm that their health problems were common, as well as unique, and related to individual prerequisites (eg, emotional reactions). Furthermore, participants emphasized the nature of information for showing people that they are not alone and for bringing hope. Those who had lived with their condition for several years offered suggestions to encourage people to do something fun, to exercise, to have goals that are important to them, and to not give up. They also mentioned the value of meeting others and the benefits of providing information about patient organizations.

The scope of information was discussed. Some users wanted a lot of information, including web links, access to video clips, and "GPS coordinates" to local services. Others emphasized that information must be brief and easy to read. Participants further identified a risk of disappointment if they found only answers to their questions without getting any solutions. Therefore, they suggested information designed to provide advice related to each question in the tool, for example, information about subsidized dental care and rehabilitation services.

Introductory Text Evokes Expectations and Facilitates Further Use

Overall, the introductory text was seen as concise, simple, and clearly stating the purpose. The layout was considered clean and appealing, with the heart-brain symbol and green leaves. After this positive first impression, participants recognized a gap between the expectations created by the introductory text and the content of the tool. One participant described that she had expected more from the tool about improving health based

on the introductory text. Notably, she had not read the advisory texts, and changed her mind when she read these texts during the interview. Several participants came up with suggestions for how the text could be revised to fill this gap.

Explanatory Texts Clarify the Question and Confirm Experiences

In general, the explanatory texts (placed underneath each question) were considered important to facilitate answering and confirm the range of related issues. Participants said that the texts clarified the questions and helped them determine whether they had appropriately interpreted them. This perception was confirmed during the interview, when one participant and her next of kin were "thinking aloud" when answering. Although some people may need help, for example, because of communication difficulties, the scope of the text was perceived as adequate. Participants generally found it valuable to read the explanatory texts, as they confirmed experiences as common and reduced potential feelings of being atypical. In pilot 1, not all participants intuitively recognized the hidden read more texts. However, when they did read them, they perceived these texts as improving their ability to give an informed answer.

In the explanatory texts, the balance between general explanations and specific examples was recognized as important. Some participants wanted more examples, whereas others felt that a general description was better. One participant mentioned that she could pay her bills but still had cognitive difficulties that interfered with her performance, for example, she was easily disturbed. Another issue mentioned was an inconsistency between the question and the provided examples. This inconsistency could be misleading for those who felt that they sufficiently managed specific activities, for example, transfer to a car, but experienced difficulties in situations demanding caution. Suggestions to improve the texts were highlighted, for example, relating to "walk and move," "pain," and "fatigue."

Advisory Texts, If Accessible, Bring Recognition and Invite Engagement

Participants appreciated the tonality of the advisory texts, embracing encouragement to engage in their own care and rehabilitation. The scope and content of the information were

considered optimal. Although some participants had previously been provided similar information, one said that seeing the information made him realize that the support he had received was adequate. They suggested naming selected support associations to enhance easy access. Despite comments that the texts were beneficial, participants also described obstacles related to accessibility. Several participants did not notice the advisory text when it was presented after submitting their answers, as they had a feeling of being “done.” It was suggested that it would be helpful to include an introduction encouraging people to read the advisory text. Participants also recommended rewording the title to include what “you can do yourself” and “how you can get support.” Suggestions to improve accessibility included placing a web link before answering or in more direct connection to each question. Additionally, the text layout was perceived as important. Those who saw all text in black on one page and compared it with the link to a text divided into sections with the specific “Strokehälsa layout” preferred the latter.

Utility and Complexity of Answering

Overview

This theme comprises the complexity of answering and includes perceptions of the questions, answer options, and interpretation of answers.

Relevance and Comprehension of the Questions

Overall, the health problems included were considered relevant and to encompass a broad range of topics. However, it was reported that some areas could be missed, for example, fatigue, vision, and swallowing. Some felt that the number of questions was just right, but others thought of additional questions to include. Answering questions not relevant to oneself was not considered a problem. In general, the questions were considered easy to comprehend. However, many participants identified the risk that if a person answered “no” to the overall question regarding “activities of daily living,” they would miss the follow-up questions, as constructed in pilot 1. Participants expressed concerns that some headings were hard to understand, particularly secondary prevention. By contrast, others reported that headings, such as cognition and spasticity, contributed to their learning.

Challenges in Answering and Interpretation

Participants discussed potential answer alternatives, such as number of options, grading, and the use of free text. For most participants, answering yes or no in combination with the related explanatory text worked out well. However, participants still expressed their wish to explain their answers—for example, that “yes” means that they are slower to do things. Participants described experiences of frustration when answering, particularly when they only experienced minor health problems. One woman described frustration when she answered “yes” to “Activities of daily living,” but only meant that she had trouble tying her shoelaces. Some argued that multiple answer options would make it easier. By contrast, others saw a risk of complicating things, especially participants with communication difficulties.

Proposals to use free text rather than yes/no were met with counterarguments. Participants who had unsatisfactory

experiences using yes and no options said that a combined solution would have improved answering. In later pilot versions, when “other concerns” and free-text boxes were added, participants valued the opportunity to raise additional issues and provide individual comments. Apart from the potential limitations of yes and no, participants acknowledged that these options made answering quick and easy. Other suggestions included using a consistent approach to subsequent questions, and the possibility of having the questions read with sound. Participants thought that the pilot versions incorporated into the platform generally functioned well. However, some failed to submit or thought that it was not easy enough to change their answers, and thus offered suggestions to improve navigability.

Interpretation of the answers was discussed. Some participants were concerned that yes and no options may not provide sufficient information to health professionals. They also felt that the phrase “more difficult after your stroke” could be challenging when they experienced a health problem as more problematic but not more difficult. In particular, participants with minor impairments reported feeling that they did not want to exaggerate the problem. They further described that their abilities were likely to shift over time or to be situation dependent, such that the answers were not unambiguous. Next of kin highlighted that the user’s perception can be opposite to that of his/her relative, indicating different views of the situation, especially several years after the stroke. The tool was viewed as a rough measure.

Capturing Needs and Value of the Tool

Overview

This theme highlights the tool’s potential to identify health problems, and the importance of dialog with a health care professional at the care visit.

Identification and Communication of One’s Own Needs

Most participants expressed that having access to the information and identifying health problems through the tool would have supported them in communicating their needs. Increased knowledge about available support was described as essential for being able to act and seek help. One participant said it would enable people to drive their case forward. Participants recognized that they often forget to bring up issues and acknowledged the benefits of making one’s health problems visible and easier to explain. Some described that insight into health problems as part of the disease picture would have encouraged them to ask the health professional questions, for example, regarding incontinence and sexuality. By contrast, one participant reflected that the absence of health problems in the tool could lead to a patient not associating this problem with his/her condition, and thus to be less likely to discuss it. Some described experiences of facing new problems after some time, for example, fatigue and return to work, and regarded the tool as helpful in this context. People with communication difficulties and 1 next of kin thought that a care visit could be enhanced by using the tool beforehand. Nevertheless, it was recognized that not all people can use the tool as an aid to identify their needs and that some would prefer a paper version.

Dialog and Making Use of the Answers

Participants strongly emphasized that use of the tool should be followed by a dialog between the health care professional and the patient. Many described a desire to explain what they meant by their answers. One participant expressed that it was sufficient to answer “yes” in the tool if the “yes” can be elaborated on in a subsequent conversation at the care visit. Another participant felt that meeting with people was preferable, compared with free text or ticking a box; however, he could see value of the combination of both. Some suggested that the use of their answers at the care visit was a fundamental prerequisite for the usefulness of the tool. Although participants considered the answers useful for health care professionals, there were concerns regarding whether they would have the necessary staff resources to fully implement this new digital service and change their ordinary routines.

Discussion

Principal Findings and Comparison With Prior Work

A participatory co-design was used to develop a digital previsit health care tool based on experiences of people with stroke, health professionals, and researchers. Integration of health information, validated questions, and digital functionality contributed to the development of a tool perceived as easy to use. The findings suggest that a condition-specific tool can confirm commonly perceived experiences and give targeted support and that the elements in the tool can be adapted to other health conditions.

This study included discussions about the optimal scope of information and questions in the tool. The theme *Valuable accessible information* describes needs that can potentially be met by eHealth services [13]. Besides condition-specific information, participants in this study particularly valued information encouraging them to be involved in their care. Accordingly, the information in the tool was created with the goals of preparing patients for active participation and enhancing their involvement in decision making [1]. Nonetheless, given the various levels of eHealth literacy [18] and the different expectations among participants, it was challenging to provide information at an optimal level for the group.

During the conceptualization phase, participants valued reliable information of various kinds, and the amount of information to be included was not yet determined. Of note, on the secure platform, the possibility of including and layering a larger amount of information was restricted, so that only brief information was included. Furthermore, the validated questions in the tool were perceived to cover the most important topics. These findings reinforce those of a previous study suggesting that the PSC directly or indirectly covers most problems [36]. The risk of health problems (eg, nutrition, sex life, and fatigue) not being covered has been previously discussed [31,37] and was mentioned in this study. A previously reported digital tool developed for long-term conditions included a more comprehensive list of nonvalidated items [17], but its usability in a clinical context remains to be tested.

It is important to reflect on whether adding more questions or measurements counteracts the perceived usability of a tool [24]. When an adapted version of the PSC was employed in combination with other measurements in a digital platform, it was used by only 11.8% of the patients, although they were offered training to use it [26]. However, usability aspects are not provided in detail. In our study, instead of increasing the number of questions, the principal decision was made to adjust the texts and encourage users to use the separate free-text option at the end, when appropriate. The hope was that the design of the tool and mode of information provision would accommodate a large group of people with long-term conditions, among whom eHealth literacy can be low [18]. Evaluation in clinical practice is important to explore whether the scope and level are optimal. Nevertheless, the level of information and questions in version 1.0 were considered a good starting point for empowering patients to be actively involved in their conversations with health care professionals.

Within the theme *Utility and complexity of answering*, aspects of reporting perceived needs and health problems were highlighted. Previous findings indicate that identifying unmet needs through self-report is complex for people [38]. Unmet needs are influenced by perceptions and experiences, such as value of independence or insights regarding available services. Participants in this study stated that their abilities were likely to shift over time and because of changing circumstances, which is consistent with previous research [38]. Moreover, people may not be capable of fully understanding and answering the questions in relation to their own situation [9,18].

In this study, a combined solution was used to facilitate answering for a broad range of patients. First, it was considered best to provide the answer options of yes/no/“choose not to answer” together with a free-text option at the end. People seem to take longer to consider the information in a question when a yes/no format is used compared with ticking a box in a list [39]. This format encourages people to think about the question in relation to their own situation, thereby preparing them for shared decision making [1]. Second, the explanatory texts underneath each question were revised as an additional solution. When using the questions at a care visit, the related dialog has been highlighted as important for ensuring that health problems are identified [31]. Similarly, participants described the explanatory texts as an asset when answering, although they did not consider the texts to replace the dialog at a visit. Third, answering was improved by clarifying that patients could explain their answers at the care visit. Overall, it is likely that the complexity of answering was decreased through the combined solutions, including the questions, answer alternatives, and the information in the tool.

Throughout the development process, multifaceted solutions were applied to accommodate limitations revealed in the interviews and to improve usability. The view of shared decision making as a holistic process, including visit preparation and the visit itself, is congruent with the service design approach [19]. Solutions to complex problems can be better solved holistically; for example, a digital tool used as part of a service [20], compared with just fragmented text presented out of its context.

In this study, the true value of the tool was perceived to depend on whether identified health problems would be addressed in the conversation at the care visit. Self-reported measures completed beforehand and received by the provider lead to patients more commonly discussing nonspecific long-term health problems, without prolonging the care visit [40]. However, using checklists [41] and previsit tools alone may not result in benefits for patients [42]. Several components are important for the delivery of effective care, such as infrastructure, people resources [43], and health care professionals' skills and motivation to provide a person-centered conversation [1,41]. Moreover, successful implementation requires consideration of the meaning value for all users, and how the team can use the tool to change their routines and improve services [44]. However, participants in this study suggested that a tool that included information could empower people to act on their own more readily and seek support.

In our study, the users' needs were in focus, in line with the service design procedure [20], and in contrast to being restricted by an existing digital system. In recent years, digital maturity in the population has increased [3], and digital health systems have changed dramatically. Therefore, attention has focused on developing a flexible and sustainable solution [28]. Only later in the development process were the core functions in the pilots transferred to a health platform that both patients and providers know and trust. The intention was for the tool to be easy to copy, modify, and connect to other platforms and contexts. Usability for the individual was considered high, as people could go through the tool quickly and found it appropriate in relation to their needs. This easy access means that the tool is more likely to be used [20,24]. Nevertheless, to accommodate patients with low eHealth literacy [10,18], a paper-based version in various languages will be developed. Together with guidelines suggesting digitally based information and support [7], the tool Strokehälsa could contribute to a move toward a more proactive health care team and patient preparedness [4]. Participants' responses supported the value of the tool and its potential to capture their needs, but both need to be tested in a larger sample in a clinical setting.

Limitations

A strength of this study is the comprehensive participatory approach, including mixed methods, enabling a deep understanding of user experiences. However, some limitations must be addressed. First, despite purposive sampling, selection bias cannot be excluded. Most participants were independent in activities of daily living and used the internet daily. Furthermore, because of COVID-19, people had to connect to the tool digitally on their own device and participate in a phone interview, which may have limited the recruitment of people

less familiar with digital tools in the amendment phase. However, support from next of kin was allowed, and the early focus groups were conducted face to face. Of note, the remote data collection worked out better than expected and yielded rich data. Altogether, the use of several sources allowed triangulation of data and a broad range of participants from different settings. The detailed descriptions of participant data analysis strengthen the transferability of the findings to other contexts.

Second, in line with the qualitative approach, attempts were made to sustain rigor and reflexivity, and interpretations were influenced by prior knowledge, for example, about stroke, the PSC, and person-centered care. The members of the research team were part of the co-design process, emphasizing the collective creativity with all stakeholders [23]. Although involvement of the researcher is part of the method, it cannot be excluded that a researcher role could have affected participant statements. However, they were encouraged to speak freely and contribute to improvements in the tool. Suggestions from participants, the expert panel, and the research team were systematically registered along with reflections in memos [35]. If controversies arose during decisions, advice was sought from members of the expert panel. Third, from a co-design perspective, the level of partnership in the study can be discussed. The patient partner was not directly involved in the initial workshop built on previous research. Although interaction on equal terms is the goal, it may not be realistic or possible for the same individual to be involved in all stages, for example, because of cognitive impairment or fatigue [29]. Nevertheless, the patient partner was continuously involved in the co-design and research process. Additionally, feedback was obtained from people with long-term conditions, health professionals providing care, and researchers in different fields. The participatory approach through service design principles led to the creation of a tool based on user needs (updated versions of the tool can be found on a webpage [45]).

Conclusions

The development process with a participatory approach resulted in a previsit digital health care tool viewed as useful for people with stroke. In this process, the integration of health information, validated questions, and digital functionality was essential to overcoming the complexity of responding to the tool's questions. Even when questions were perceived as easy to comprehend, the additional information supported answering and confirmed patients' experiences. Moreover, the information encouraged people to develop their answers in dialog with the health care professional. However, larger studies that include evaluation in conjunction with a clinical visit are needed. The tool is freely shared to be adapted and improved in different contexts for ecological validity.

Acknowledgments

We thank people with stroke for their participation and their willingness to share their experiences. We also thank the health care professionals and researchers on the expert panel for their valuable input to improve the tool. Special thanks to the service designer Michael Wide for his work with prototyping and to the information technology consultant Robin Björk for his work with adapting the tool to the secure platform. The authors thank Dr. Kate Bramley-Moore for translation help regarding the quotations. This study was funded in part by grants from the Swedish Science Council (VR2017-00946), the Swedish state under an agreement

between the Swedish government and the county councils, the ALF agreement (ALFGBG-718711, ALFGBG-943103); the Swedish state under the agreement of the Swedish government and the county councils; Heart and Lung Foundation; the Swedish Brain Foundation; the Local Research and Development Board for Gothenburg and Södra Bohuslän (VGFOUGSB-76712); Sahlgrenska Hospitals Foundation (SU-892141, SU-793481); the Swedish Stroke Association; an unconditional grant from Allergan; the Hjalmar Svenssons Foundation; Göteborg Foundation for Neurological Research; Rune and Ulla Amlövs Foundation; Per-Olof Ahls Foundation; John and Brit Wennerströms Foundation; and the Greta and Einar Askers Foundation.

Conflicts of Interest

None declared.

Multimedia Appendix 1

The question guide used in focus groups and individual interviews. (a: Only in the focus groups; b: Only in the individual interviews).

[[DOCX File, 13 KB - humanfactors_v9i2e35478_app1.docx](#)]

Multimedia Appendix 2

Description of amendments and revisions in the pilot versions. a: PSC= post-stroke checklist, b: ADL= activities of daily living.

[[DOCX File, 16 KB - humanfactors_v9i2e35478_app2.docx](#)]

Multimedia Appendix 3

Screenshots of the first prototype and version 1.0 of the digital tool Strokehälsa.

[[PDF File \(Adobe PDF File\), 4901 KB - humanfactors_v9i2e35478_app3.pdf](#)]

References

1. Ekman I, Swedberg K, Taft C, Lindseth A, Norberg A, Brink E, et al. Person-centered care--ready for prime time. *Eur J Cardiovasc Nurs* 2011 Dec;10(4):248-251. [doi: [10.1016/j.ejcnurse.2011.06.008](#)] [Medline: [21764386](#)]
2. Granström E, Wannheden C, Brommels M, Hvitfeldt H, Nyström ME. Digital tools as promoters for person-centered care practices in chronic care? Healthcare professionals' experiences from rheumatology care. *BMC Health Serv Res* 2020 Dec 01;20(1):1108 [FREE Full text] [doi: [10.1186/s12913-020-05945-5](#)] [Medline: [33261602](#)]
3. The Digital Economy and Society Index (DESI). European Commission. 2021 24. URL: <https://digital-strategy.ec.europa.eu/en/policies/desi-sweden> [accessed 2022-05-23]
4. Epping-Jordan JE, Pruitt SD, Bengoa R, Wagner EH. Improving the quality of health care for chronic conditions. *Qual Saf Health Care* 2004 Aug;13(4):299-305 [FREE Full text] [doi: [10.1136/qhc.13.4.299](#)] [Medline: [15289634](#)]
5. Feigin VL, Norrving B, Mensah GA. Global Burden of Stroke. *Circ Res* 2017 Feb 03;120(3):439-448. [doi: [10.1161/CIRCRESAHA.116.308413](#)] [Medline: [28154096](#)]
6. GBD 2019 Stroke Collaborators. Global, regional, and national burden of stroke and its risk factors, 1990-2019: a systematic analysis for the Global Burden of Disease Study 2019. *Lancet Neurol* 2021 Oct;20(10):795-820 [FREE Full text] [doi: [10.1016/S1474-4422\(21\)00252-0](#)] [Medline: [34487721](#)]
7. Norrving B, Barrick J, Davalos A, Dichgans M, Cordonnier C, Guekht A, et al. Action Plan for Stroke in Europe 2018-2030. *Eur Stroke J* 2018 Dec;3(4):309-336 [FREE Full text] [doi: [10.1177/2396987318808719](#)] [Medline: [31236480](#)]
8. Pindus DM, Mullis R, Lim L, Wellwood I, Rundell AV, Abd Aziz NA, et al. Stroke survivors' and informal caregivers' experiences of primary care and community healthcare services - A systematic review and meta-ethnography. *PLoS One* 2018;13(2):e0192533 [FREE Full text] [doi: [10.1371/journal.pone.0192533](#)] [Medline: [29466383](#)]
9. Friis K, Lasgaard M, Osborne RH, Maindal HT. Gaps in understanding health and engagement with healthcare providers across common long-term conditions: a population survey of health literacy in 29,473 Danish citizens. *BMJ Open* 2016 Jan 14;6(1):e009627 [FREE Full text] [doi: [10.1136/bmjopen-2015-009627](#)] [Medline: [26769783](#)]
10. Cheng C, Beauchamp A, Elsworth GR, Osborne RH. Applying the electronic health literacy lens: systematic review of electronic health interventions targeted at socially disadvantaged groups. *J Med Internet Res* 2020 Aug 13;22(8):e18476 [FREE Full text] [doi: [10.2196/18476](#)] [Medline: [32788144](#)]
11. Antonenko K, Paciaroni M, Sokolova L, Pezzella FR. Digital health in stroke medicine: what are the opportunities for stroke patients? *Curr Opin Neurol* 2021 Feb 01;34(1):27-37. [doi: [10.1097/WCO.0000000000000891](#)] [Medline: [33278145](#)]
12. Healthcare guide 1177. 1177 Vårdguiden. 2022. URL: <https://www.1177.se/en/Vastra-Gotaland/other-languages/other-languages/About-1177.se/about-1177-varldguiden> [accessed 2022-05-23]
13. Davoody N, Koch S, Krakau I, Hägglund M. Post-discharge stroke patients' information needs as input to proposing patient-centred eHealth services. *BMC Med Inform Decis Mak* 2016 Jun 07;16:66 [FREE Full text] [doi: [10.1186/s12911-016-0307-2](#)] [Medline: [27267889](#)]

14. Stacey D, Légaré F, Lewis K, Barry MJ, Bennett CL, Eden KB, et al. Decision aids for people facing health treatment or screening decisions. *Cochrane Database Syst Rev* 2017 Apr 12;4:CD001431 [FREE Full text] [doi: [10.1002/14651858.CD001431.pub5](https://doi.org/10.1002/14651858.CD001431.pub5)] [Medline: [28402085](https://pubmed.ncbi.nlm.nih.gov/28402085/)]
15. Qudah B, Luetsch K. The influence of mobile health applications on patient - healthcare provider relationships: A systematic, narrative review. *Patient Educ Couns* 2019 Jun;102(6):1080-1089. [doi: [10.1016/j.pec.2019.01.021](https://doi.org/10.1016/j.pec.2019.01.021)] [Medline: [30745178](https://pubmed.ncbi.nlm.nih.gov/30745178/)]
16. Holt JM, Cusatis R, Winn A, Asan O, Spanbauer C, Williams JS, et al. Impact of pre-visit contextual data collection on patient-physician communication and patient activation: a randomized trial. *J Gen Intern Med* 2021 Nov;36(11):3321-3329. [doi: [10.1007/s11606-020-06583-7](https://doi.org/10.1007/s11606-020-06583-7)] [Medline: [33559067](https://pubmed.ncbi.nlm.nih.gov/33559067/)]
17. Seljelid B, Varsi C, Solberg Nes L, Stenehjem A, Bollerslev J, Børøsund E. Content and system development of a digital patient-provider communication tool to support shared decision making in chronic health care: InvolveMe. *BMC Med Inform Decis Mak* 2020 Mar 04;20(1):46 [FREE Full text] [doi: [10.1186/s12911-020-1065-8](https://doi.org/10.1186/s12911-020-1065-8)] [Medline: [32131808](https://pubmed.ncbi.nlm.nih.gov/32131808/)]
18. Neter E, Brainin E. eHealth literacy: extending the digital divide to the realm of health information. *J Med Internet Res* 2012 Jan 27;14(1):e19 [FREE Full text] [doi: [10.2196/jmir.1619](https://doi.org/10.2196/jmir.1619)] [Medline: [22357448](https://pubmed.ncbi.nlm.nih.gov/22357448/)]
19. Griffioen I, Melles M, Stiggelbout A, Snelders D. The potential of service design for improving the implementation of shared decision-making. *Design for Health* 2017 Oct 24;1(2):194-209. [doi: [10.1080/24735132.2017.1386944](https://doi.org/10.1080/24735132.2017.1386944)]
20. Stickdorn M, Hormess ME. *This Is Service Design Doing: Applying Service Design Thinking in the Real World*. Farnham, QC, Canada: O'Reilly Media; 2018.
21. Göttgens I, Oertelt-Prigione S. The application of human-centered design approaches in health research and innovation: a narrative review of current practices. *JMIR Mhealth Uhealth* 2021 Dec 06;9(12):e28102 [FREE Full text] [doi: [10.2196/28102](https://doi.org/10.2196/28102)] [Medline: [34874893](https://pubmed.ncbi.nlm.nih.gov/34874893/)]
22. Steen M, Manschot M, De Koning N. Benefits of co-design in service design projects. *International Journal of Design* 2011;5(2):53-60 [FREE Full text]
23. Sanders EB, Stappers PJ. Co-creation and the new landscapes of design. *CoDesign* 2008 Mar;4(1):5-18. [doi: [10.1080/15710880701875068](https://doi.org/10.1080/15710880701875068)]
24. Davis FD. Perceived usefulness, perceived ease of use, and user acceptance of information technology. *MIS Quarterly* 1989 Sep;13(3):319. [doi: [10.2307/249008](https://doi.org/10.2307/249008)]
25. Giroux D, Tremblay M, Latulippe K, Provencher V, Poulin V, Giguere A, et al. Promoting identification and use of aid resources by caregivers of seniors: co-design of an electronic health tool. *JMIR Aging* 2019 Aug 22;2(2):e12314 [FREE Full text] [doi: [10.2196/12314](https://doi.org/10.2196/12314)] [Medline: [31518284](https://pubmed.ncbi.nlm.nih.gov/31518284/)]
26. Willeit P, Toell T, Boehme C, Krebs S, Mayer L, Lang C, STROKE-CARD study group. STROKE-CARD care to prevent cardiovascular events and improve quality of life after acute ischaemic stroke or TIA: A randomised clinical trial. *EClinicalMedicine* 2020 Aug;25:100476 [FREE Full text] [doi: [10.1016/j.eclinm.2020.100476](https://doi.org/10.1016/j.eclinm.2020.100476)] [Medline: [32954239](https://pubmed.ncbi.nlm.nih.gov/32954239/)]
27. Michalski D, Prost A, Handel T, Schreiber M, Tylcz J, Geisler D, et al. The PostStroke-Manager - combining mobile, digital and sensor-based technology with personal assistance: protocol of the feasibility study. *Neurol Res Pract* 2021 Sep 09;3(1):53 [FREE Full text] [doi: [10.1186/s42466-021-00137-w](https://doi.org/10.1186/s42466-021-00137-w)] [Medline: [34503568](https://pubmed.ncbi.nlm.nih.gov/34503568/)]
28. Morgan D. *Integrating Qualitative and Quantitative Methods: A Pragmatic Approach*. London, UK: Sage Publications; 2014.
29. Cowdell F, Dyson J, Sykes M, Dam R, Pendleton R. How and how well have older people been engaged in healthcare intervention design, development or delivery using co-methodologies: A scoping review with narrative summary. *Health Soc Care Community* 2022 Feb;30(2):776-798. [doi: [10.1111/hsc.13199](https://doi.org/10.1111/hsc.13199)] [Medline: [33103313](https://pubmed.ncbi.nlm.nih.gov/33103313/)]
30. Malterud K, Siersma VD, Guassora AD. Sample size in qualitative interview studies: guided by information power. *Qual Health Res* 2016 Nov;26(13):1753-1760. [doi: [10.1177/1049732315617444](https://doi.org/10.1177/1049732315617444)] [Medline: [26613970](https://pubmed.ncbi.nlm.nih.gov/26613970/)]
31. Kjörk EK, Carlsson G, Sunnerhagen KS, Lundgren-Nilsson Å. Experiences using the poststroke checklist in Sweden with a focus on feasibility and relevance: a mixed-method design. *BMJ Open* 2019 May 09;9(5):e028218 [FREE Full text] [doi: [10.1136/bmjopen-2018-028218](https://doi.org/10.1136/bmjopen-2018-028218)] [Medline: [31072862](https://pubmed.ncbi.nlm.nih.gov/31072862/)]
32. Philp I, Brainin M, Walker MF, Ward AB, Gillard P, Shields AL, Global Stroke Community Advisory Panel. Development of a poststroke checklist to standardize follow-up care for stroke survivors. *J Stroke Cerebrovasc Dis* 2013 Oct;22(7):e173-e180 [FREE Full text] [doi: [10.1016/j.jstrokecerebrovasdis.2012.10.016](https://doi.org/10.1016/j.jstrokecerebrovasdis.2012.10.016)] [Medline: [23265778](https://pubmed.ncbi.nlm.nih.gov/23265778/)]
33. Krueger RA. *Focus Groups: A Practical Guide for Applied Research*. Thousand Oaks, CA: Sage Publications; 2014.
34. Beatty P, Willis G. Research synthesis: the practice of cognitive interviewing. *Public Opinion Quarterly* 2007 Jun 05;71(2):287-311. [doi: [10.1093/poq/nfm006](https://doi.org/10.1093/poq/nfm006)]
35. Ritchie J, Lewis J, Nicholls C, Ormston R. *Qualitative research practice: a guide for social science students and researchers*. Thousand Oaks, CA: Sage Publications; 2013.
36. Ward A, Chen C, Norrving B, Gillard P, Walker M, Blackburn S, Global Stroke Community Advisory Panel (GSCAP). Evaluation of the Post Stroke Checklist: a pilot study in the United Kingdom and Singapore. *Int J Stroke* 2014 Oct;9 Suppl A100:76-84 [FREE Full text] [doi: [10.1111/ijvs.12291](https://doi.org/10.1111/ijvs.12291)] [Medline: [25088427](https://pubmed.ncbi.nlm.nih.gov/25088427/)]
37. Turner GM, Mullis R, Lim L, Kreit L, Mant J. Using a checklist to facilitate management of long-term care needs after stroke: insights from focus groups and a feasibility study. *BMC Fam Pract* 2019 Jan 04;20(1):2 [FREE Full text] [doi: [10.1186/s12875-018-0894-3](https://doi.org/10.1186/s12875-018-0894-3)] [Medline: [30609920](https://pubmed.ncbi.nlm.nih.gov/30609920/)]

38. Shannon RL, Forster A, Hawkins RJ. A qualitative exploration of self-reported unmet need one year after stroke. *Disabil Rehabil* 2016 Oct;38(20):2000-2007. [doi: [10.3109/09638288.2015.1107784](https://doi.org/10.3109/09638288.2015.1107784)] [Medline: [26733052](https://pubmed.ncbi.nlm.nih.gov/26733052/)]
39. Smyth JD. Comparing check-all and forced-choice question formats in web surveys. *Public Opinion Quarterly* 2006 Mar 01;70(1):66-77. [doi: [10.1093/poq/nfj007](https://doi.org/10.1093/poq/nfj007)]
40. Velikova G, Booth L, Smith AB, Brown PM, Lynch P, Brown JM, et al. Measuring quality of life in routine oncology practice improves communication and patient well-being: a randomized controlled trial. *J Clin Oncol* 2004 Feb 15;22(4):714-724. [doi: [10.1200/JCO.2004.06.078](https://doi.org/10.1200/JCO.2004.06.078)] [Medline: [14966096](https://pubmed.ncbi.nlm.nih.gov/14966096/)]
41. Morrissey M, Shepherd E, Kinley E, McClatchey K, Pinnock H. Effectiveness and perceptions of using templates in long-term condition reviews: a systematic synthesis of quantitative and qualitative studies. *Br J Gen Pract* 2021 Sep;71(710):e652-e659 [FREE Full text] [doi: [10.3399/BJGP.2020.0963](https://doi.org/10.3399/BJGP.2020.0963)] [Medline: [33690148](https://pubmed.ncbi.nlm.nih.gov/33690148/)]
42. Kinnersley P, Edwards A, Hood K, Ryan R, Prout H, Cadbury N, et al. Interventions before consultations to help patients address their information needs by encouraging question asking: systematic review. *BMJ* 2008 Jul 16;337:a485 [FREE Full text] [doi: [10.1136/bmj.a485](https://doi.org/10.1136/bmj.a485)] [Medline: [18632672](https://pubmed.ncbi.nlm.nih.gov/18632672/)]
43. Bernhardt J, Urimubenshi G, Gandhi DBC, Eng JJ. Stroke rehabilitation in low-income and middle-income countries: a call to action. *Lancet* 2020 Oct 31;396(10260):1452-1462. [doi: [10.1016/S0140-6736\(20\)31313-1](https://doi.org/10.1016/S0140-6736(20)31313-1)] [Medline: [33129396](https://pubmed.ncbi.nlm.nih.gov/33129396/)]
44. Shaw J, Agarwal P, Desveaux L, Palma DC, Stamenova V, Jamieson T, et al. Beyond "implementation": digital health innovation and service design. *NPJ Digit Med* 2018;1:48 [FREE Full text] [doi: [10.1038/s41746-018-0059-8](https://doi.org/10.1038/s41746-018-0059-8)] [Medline: [31304327](https://pubmed.ncbi.nlm.nih.gov/31304327/)]
45. The digital tool Strokehälsa: information and user instructions. Institute of Neuroscience and Physiology. Gothenburg, Sweden: University of Gothenburg; 2022. URL: <https://www.gu.se/en/neuroscience-physiology/strokehalsatm> [accessed 2022-05-23]

Abbreviations

PSC: Post-Stroke Checklist

Strokehälsa: Strokehealth

TAM: Technological Acceptance Model

Edited by A Kushniruk; submitted 18.12.21; peer-reviewed by E Børøsund, C Wannheden, E Neter, M Fiordelli; comments to author 06.03.22; revised version received 29.03.22; accepted 01.05.22; published 03.06.22.

Please cite as:

Kjörk EK, Sunnerhagen KS, Lundgren-Nilsson Å, Andersson AK, Carlsson G

Development of a Digital Tool for People With a Long-Term Condition Using Stroke as a Case Example: Participatory Design Approach

JMIR Hum Factors 2022;9(2):e35478

URL: <https://humanfactors.jmir.org/2022/2/e35478>

doi: [10.2196/35478](https://doi.org/10.2196/35478)

PMID: [35657650](https://pubmed.ncbi.nlm.nih.gov/35657650/)

©Emma K Kjörk, Katharina S Sunnerhagen, Åsa Lundgren-Nilsson, Anders K Andersson, Gunnel Carlsson. Originally published in *JMIR Human Factors* (<https://humanfactors.jmir.org>), 03.06.2022. This is an open-access article distributed under the terms of the Creative Commons Attribution License (<https://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in *JMIR Human Factors*, is properly cited. The complete bibliographic information, a link to the original publication on <https://humanfactors.jmir.org>, as well as this copyright and license information must be included.

Corrigenda and Addenda

Correction: Improving Pelvic Floor Muscle Training Adherence Among Pregnant Women: Validation Study

Aida Jaffar^{1,2}, MMed; Sherina Mohd-Sidik¹, PhD; Chai Nien Foo³, PhD; Novia Admodisaastro⁴, PhD; Sobihatun Nur Abdul Salam⁵, PhD; Noor Diana Ismail⁶, MMed

¹Department of Psychiatry, Faculty of Medicine and Health Sciences, Universiti Putra Malaysia, Selangor, Malaysia

²Primary Care Unit, Faculty of Medicine and Defence Health, Universiti Pertahanan Nasional Malaysia, Wilayah Persekutuan, Malaysia

³Department of Population Medicine, Universiti Tunku Abdul Rahman, Selangor, Malaysia

⁴Software Engineering & Information System Department, Faculty of Computer Science & Information Technology, Universiti Putra Malaysia, Selangor, Malaysia

⁵School of Multimedia Technology and Communication, College of Arts and Sciences, Universiti Utara Malaysia, Kedah, Malaysia

⁶Klinik Kesihatan Bt 9 Cheras, Ministry of Health, Selangor, Malaysia

Corresponding Author:

Sherina Mohd-Sidik, PhD

Department of Psychiatry

Faculty of Medicine and Health Sciences

Universiti Putra Malaysia

Serdang

Selangor, 43400

Malaysia

Phone: 60 3 9769 2541

Fax: 60 3 9769 2706

Email: sherina@upm.edu.my

Related Article:

Correction of: <https://humanfactors.jmir.org/2022/1/e30989>

(*JMIR Hum Factors* 2022;9(2):e38175) doi:[10.2196/38175](https://doi.org/10.2196/38175)

In “Improving Pelvic Floor Muscle Training Adherence Among Pregnant Women: Validation Study” (*JMIR Hum Factors* 2022;9(1):e30989), the following errors were noted.

1. Abstract:

In the originally published paper, the first sentence in the abstract was stated as

Mobile health apps, for example, the Tüt, have been shown to be potentially effective in improving pelvic floor muscle training (PFMT) among women, but their effectiveness in pregnant women was limited.

This has been corrected to:

Mobile health apps, for example, the Tüt, have been shown to be potentially effective in improving pelvic floor muscle training (PFMT) among women, but they have not yet been studied among pregnant women.

2. Methods, Intervention Mapping:

The originally published paper was missing two references for this statement:

The outcomes of the intention are self-efficacy (17 questions) and adherence (6 questions).

This has been corrected to:

The outcomes of the intention are self-efficacy (17 questions) (41) and adherence (6 questions) (42).

3. Methods, Cross-Sectional Study:

The originally published paper was missing two references for this statement:

The findings from this study provided input for the content of their educational videos and short notes on PFMT, which were captured as frequently asked questions (FAQ).

This has been corrected to:

The findings from this study provided input for the content of their educational videos and short notes on PFMT (45, 46) which were captured as frequently asked questions (FAQ).

4. Results:

The originally published paper stated the following in row 1, column 2 of Table 5:

1.System credibility-expertise and authority.

2. Primary support-Virtual rehearsal principle

This has been corrected to:

System credibility-expertise and authority

5. Results:

The originally published paper stated the following as the title for the first column of Table 5:

COM-B model and behavioral change techniques incorporated in the mHealth app.

This has been corrected to:

COM-B model and features of the mHealth app.

6. Discussion:

The originally published paper was missing one reference for this statement:

The PSD component of the system's credibility and trustworthiness, with the expertise involved in the development, may add to the user's sense of safety and reliability regarding the KEPT app.

This has been corrected to:

The PSD component of the system's credibility and trustworthiness (55), with the expertise involved in the development, may add to the user's sense of safety and reliability regarding the KEPT app.

6. References:

In the corrected paper, the following citations have been newly added to the Reference List. As these new references have been numbered per the order of their in-text citations, the remaining citations in the reference list have been renumbered accordingly.

41. Sacomori C, Cardoso FL, Porto IP, Negri NB. The development and psychometric evaluation of a self-efficacy scale for practicing pelvic floor exercises. *Brazilian Journal of Physical Therapy*; 2013. [doi: 10.1590/S1413-35552012005000104]

42. Newman-Beinart NA, Norton S, Dowling D, Gavriloff D, Vari C, Weinman JA, Godfrey EL. The development and initial psychometric evaluation of a measure assessing adherence to prescribed exercise: the Exercise Adherence Rating Scale (EARS). *Physiotherapy*, 103(2); 2017, 180–185. [doi: 10.1016/j.physio.2016.11.001]

45. Alagirisamy P, Mohd Sidik S. *Pelvic Floor Muscle Exercises During and After Pregnancy*. Universiti Putra Malaysia Press Serdang 2020.

46. Bo K, Berghmans B, Morkved S, Van Kampen M. *Evidence-Based Physical Therapy for the Pelvic Floor-E-Book: Bridging Science and Clinical Practice*, 2nd ed. London, UK: Elsevier Health Sciences; 2014. 3.

55. Asklund I, Nyström E, Sjöström M, Umefjord G, Stenlund H, Samuelsson E. Mobile app for treatment of stress urinary incontinence: A randomized controlled trial. *Neurourol Urodyn* 2017 Jun;36(5):1369-1376. [doi: 10.1002/nau.23116] [Medline: 27611958]

The correction will appear in the online version of the paper on the JMIR Publications website on April 11, 2022, together with the publication of this correction notice. Because this was made after submission to PubMed, PubMed Central, and other full-text repositories, the corrected article has also been resubmitted to those repositories.

Submitted 21.03.22; this is a non-peer-reviewed article; accepted 23.03.22; published 11.04.22.

Please cite as:

Jaffar A, Mohd-Sidik S, Foo CN, Admodisastro N, Abdul Salam SN, Ismail ND

Correction: Improving Pelvic Floor Muscle Training Adherence Among Pregnant Women: Validation Study

JMIR Hum Factors 2022;9(2):e38175

URL: <https://humanfactors.jmir.org/2022/2/e38175>

doi: [10.2196/38175](https://doi.org/10.2196/38175)

PMID: [35404829](https://pubmed.ncbi.nlm.nih.gov/35404829/)

©Aida Jaffar, Sherina Mohd-Sidik, Chai Nien Foo, Novia Admodisastro, Sobihatun Nur Abdul Salam, Noor Diana Ismail. Originally published in *JMIR Human Factors* (<https://humanfactors.jmir.org/>), 11.04.2022. This is an open-access article distributed under the terms of the Creative Commons Attribution License (<https://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in *JMIR Human Factors*, is properly cited. The complete bibliographic information, a link to the original publication on <https://humanfactors.jmir.org/>, as well as this copyright and license information must be included.

Original Paper

Smartphone Alcohol Use Disorder Recovery Apps: Cross-sectional Survey of Behavioral Intention to Use

Rijuta Menon¹, BSc, MSc; Julien Meyer¹, BA, MSc, PhD; Pria Nippak¹, BA, BSc, MSc, PhD; Housne Begum¹, BSc, MSc, PhD

School of Health Services Management, Ted Rogers School of Management, Ryerson University, Toronto, ON, Canada

Corresponding Author:

Julien Meyer, BA, MSc, PhD

School of Health Services Management

Ted Rogers School of Management

Ryerson University

350 Victoria Street

Office: TRS 3-081

Toronto, ON, M5B 2K3

Canada

Phone: 1 4169951338

Email: julien.meyer@ryerson.ca

Abstract

Background: Alcohol use disorder (AUD) carries a huge health and economic cost to society. Effective interventions exist but numerous challenges limit their adoption, especially in a pandemic context. AUD recovery apps (AUDRA) have emerged as a potential complement to in-person interventions. They are easy to access and show promising results in terms of efficacy. However, they rely on individual adoption decisions and remain underused.

Objective: The aim of this survey study is to explore the beliefs that determine the intention to use AUDRA.

Methods: We conducted a cross-sectional survey study of people with AUD. We used the Unified Theory of Acceptance and Use of Technology, which predicts use and behavioral intention to use based on performance expectancy, effort expectancy, social influence, and facilitating conditions. Participants were recruited directly from 2 sources; first, respondents at addiction treatment facilities in Ontario, Canada, were contacted in person, and they filled a paper form; second, members from AUD recovery support groups on social media were contacted and invited to fill an internet-based survey. The survey was conducted between October 2019 and June 2020.

Results: The final sample comprised 159 participants (124 involved in the web-based survey and 35 in the paper-based survey) self-identifying somewhat or very much with AUD. Most participants ($n=136$, 85.5%) were aware of AUDRA and those participants scored higher on performance expectancy, effort expectancy, and social influence. Overall, the model explains 35.4% of the variance in the behavioral intention to use AUDRA and 11.1% of the variance in use. Social influence ($P=.31$), especially for women ($P=.23$), and effort expectancy ($P=.25$) were key antecedents of behavioral intention. Facilitating conditions were not significant overall but were moderated by age ($P=.23$), suggesting that it matters for older participants. Performance expectancy did not predict behavioral intention, which is unlike many other technologies but confirms other findings associated with mobile health (mHealth). Open-ended questions suggest that privacy concerns may significantly influence the use of AUDRA.

Conclusions: This study suggests that unlike many other technologies, the adoption of AUDRA is not mainly determined by utilitarian factors such as performance expectancy. Rather, effort expectancy and social influence play a key role in determining the intention to use AUDRA.

(*JMIR Hum Factors* 2022;9(2):e33493) doi:[10.2196/33493](https://doi.org/10.2196/33493)

KEYWORDS

mobile health; alcohol use disorder; disease management; mobile apps; Unified Theory of Acceptance and Use of Technology

Introduction

Alcohol causes 3.3 million deaths a year worldwide, close to 6% of all deaths [1]. Many of these deaths are associated with alcohol use disorder (AUD), defined as “a problematic pattern of alcohol use accompanied by clinically significant impairment or distress” [2]. Treatment and engagement with recovery activities, such as brief interventions, motivational enhancements, and cognitive behavior therapies, are integral to avoiding disease progression [3]. They are well accepted and effective. However, they usually require substantial time, money, and resources; moreover, they depend predominantly on the skill of the clinician and can be stigmatizing [3].

With the advent of smartphones, mobile health (mHealth) apps have been developed to address AUD recovery. These apps can provide information and advice on how to address the condition and help users track their behavior. They serve as accessible, widespread, cost-effective, dependable, individualized, and anonymous alternatives or complements to traditional interventions [3]. These apps have also proved invaluable in the context of the COVID-19 pandemic, which has aggravated addiction issues while severely restricting access to in-person support services. In a 2019 literature review on the efficiency of AUD recovery apps (AUDRA), 63% (n=12) of the 19 studies considered found significant evidence of positive outcomes, 32% (n=6) found none, and 5% (n=1) found negative outcomes for some users [1]. Positive outcomes included decreased alcohol consumption, decreased episodes of binge drinking and alcohol-related injuries, and decreased addiction levels. Despite these benefits, evidence from mHealth app studies indeed suggest low adoption rates [4,5], and studies about the acceptance of mental health apps particularly suggest that potential users remain unconvinced of their usefulness.

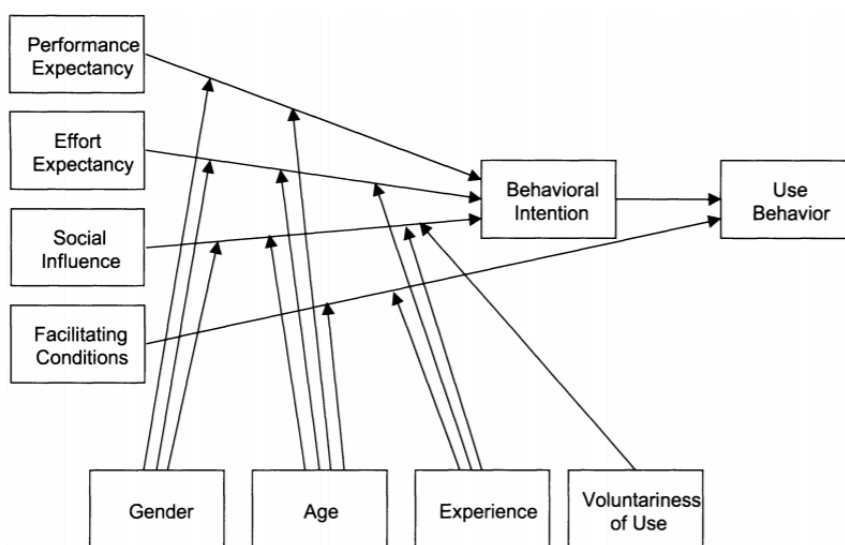
Technology adoption has been the subject of significant research attention and conceptualization. The Unified Theory of Acceptance and Use of Technology (UTAUT) is a well-established theory of acceptance of consumer technology [6]. It unifies 8 prominent and competing models of user acceptance of new information technologies [6,7]. With UTAUT2, the model was extended from organizational adoption to a consumer use context [7]. This theory is a good predictor of the intention to use mHealth [8,9], but it has not been used yet to investigate beliefs related to AUDRA. This study was designed to investigate the potential factors contributing to AUDRA adoption among people with AUD.

Methods

Study Design and Survey Instrument

This study is a cross-sectional survey of nonusers or existing users of AUDRA. The survey covered the factors contributing to the behavioral intention to use smartphone AUD recovery apps among participants (it targeted use of AUDRA in general and not of any specific app). The UTAUT framework and model questionnaire items (Figure 1 and Textbox 1) were adapted to measure the constructs, particularly its operationalizations from UTAUT2. UTAUT predicts that the behavioral intention to use a technology depends on four factors: (1) performance expectancy, defined as the degree to which using a technology will provide benefits to consumers in performing certain activities; (2) effort expectancy, defined as the degree of ease associated with consumers’ use of technology; (3) social influence, defined as the extent to which consumers perceive that important others (eg, family and friends) believe they should use a particular technology; and (4) facilitating conditions, defined as consumers’ perceptions of the resources and support available to perform a behavior [6,7].

Figure 1. Unified Theory of Acceptance and Use of Technology research model showing the complete theoretical model with the moderating relationships [6].



The constructs of hedonic motivation, price value, and habit from UTAUT2 were removed. They are not applicable to this study as AUDRA are not primarily designed for enjoyment; almost all AUDRA are free on app stores, and AUDRA are still new and rare, which diminish the importance of habit and

experience. Age and gender also moderate these relations. Figure 1 shows the theoretical model with the moderating relationships.

The constructs were measured by adapting the 16 corresponding items from UTAUT2 [7] using a 5-point Likert scale ranging

from “strongly disagree” to “strongly agree,” except for behavioral intention that had choices “yes,” “no,” or “maybe” and use, which used a 6-point Likert scale ranging from “everyday” to “at least once a year” (Textbox 1). A follow-up

survey was conducted 6 months later to investigate the subsequent usage behavior. The study was approved by Ryerson’s Research Ethics Board (approval reference number: 2019-277).

Textbox 1. Survey items used for each construct.

<p>Performance expectancy</p> <ol style="list-style-type: none"> 1. I find/would find Smartphone Alcohol Use Disorder (AUD) recovery apps useful in complementing the daily activities I do to help me recover. 2. Using Smartphone AUD recovery apps helps/would help me learn recovery skills more quickly. 3. Using Smartphone AUD recovery apps helps/would help me increase the effectiveness of activities I do to help me recover. <p>Effort expectancy</p> <ol style="list-style-type: none"> 4. Learning how to use Smartphone AUD recovery apps is/would be easy for me. 5. My interaction with Smartphone AUD recovery apps is/would be clear and understandable. 6. I find/would find Smartphone AUD recovery apps addiction recovery apps easy to use. 7. It is/would be easy for me to become skillful at using Smartphone AUD recovery apps. <p>Social influence</p> <ol style="list-style-type: none"> 8. People who are important to me think that I should use Smartphone AUD recovery apps. 9. Caregivers think that I should use Smartphone AUD recovery apps. 10. People who influence my behavior think that I should use Smartphone AUD recovery apps. 11. People whose opinions that I value prefer that I use Smartphone AUD recovery apps. <p>Facilitating conditions</p> <ol style="list-style-type: none"> 12. I have the resources necessary to use Smartphone AUD recovery apps. 13. I have the knowledge necessary to use Smartphone AUD recovery apps. 14. Smartphone AUD recovery apps are compatible with other technologies I use. 15. I can get help from others to use Smartphone AUD recovery apps. <p>Behavioral intention</p> <ol style="list-style-type: none"> 16. Do you intend to use or keep using a Smartphone AUD recovery app(s)? <p>Use</p> <ol style="list-style-type: none"> 17. If you are using a Smartphone app that assists with recovery AUD, how often do you use it?

Recruitment

Participants were aged 18 years and older, self-identified as having an AUD, and owned a smartphone. Data were collected between October 2019 and June 2020. The survey was offered to participants in 2 modalities. The first was in a pen and paper format, with participants recruited from 6 AUD treatment facilities in Ontario, Canada. Second, an internet-based version of the survey was shared on various English-speaking AUD recovery groups. Participants were offered a gift certificate for their participation. A second follow-up survey was conducted to track usage longitudinally, but it was discarded due to an insufficient response rate. In the partial least squares-structural equation model (PLS-SEM), the minimum sample size is 10 times the number of paths targeting a particular construct. In our study, this means a minimum of 40 respondents [10].

Statistical Analysis

Internal validity was evaluated using the Cronbach alpha and composite reliability (CR) [11]. Values for the Cronbach alpha and CR are considered satisfactory if they are between 0.7 and

0.9 [12]. Convergent validity was assessed using the outer loadings of the indicators and the values of the average variance extracted (AVE) [11]. To help establish convergent validity on a construct, the outer loadings should be 0.708 or higher and the AVE value must be 0.5 or higher to indicate that the construct explains more than 50% of the variance of its indicators [11].

The heterotrait-monotrait ratio was used to assess the discriminant validity between constructs. When constructs are conceptually more distinct, as is the case with the constructs of UTAUT, a lower conservative threshold of 0.85 is suggested such that values above this threshold indicate a lack of discriminant validity [11].

The results of the survey were analyzed using SPSS Statistics (version 26; IBM Corporation) and SmartPLS 3 (version 3.2.9; SmartPLS GmbH). SPSS Statistics was used for descriptive statistics and chi-square tests were performed to test the associations between variables and differences in the mean scores for variables; their determinants between the 2 groups were assessed using *t* tests at a 95% CI. As the focus of this

study was on identifying the antecedents of smartphone AUD recovery app adoption, there were no exclusion criteria in place to exempt the responses of those who did not possess prior knowledge about the existence of these apps. PLS-SEM was used to test the research model (Figure 1) for its reliability, convergent validity, and the discriminant validity of the constructs. The structural model was assessed using R^2 and bootstrapping tests were conducted to examine the statistical significance (taken at 95% CI) of the path coefficients [11]. For the PLS algorithms and bootstrapping calculations, missing data were treated with mean value replacement. SmartPLS 3 was used to test the theoretical model.

The open-ended questions aimed to determine why the participants used or did not use AUDRA. The comments were analyzed quantitatively by themes [13]. Although the low rate of response for these questions did not allow for deriving meaningful statistics, it was sufficient to identify some recurring themes.

Results

User Statistics

A total of 1792 surveys were completed. However, most web-based surveys had to be excluded, with 900 excluded for multiple participations, 416 for answering randomly or incompletely, and 317 for not meeting the inclusion criteria (not identifying with AUD or not owning a smartphone) Finally, 159 surveys (124 web-based and 35 paper surveys) could be used.

Table 1 provides the background characteristics of the respondents. The 159 respondents comprised 111 (69.8%) males, 45 (28.3%) females, and 3 (1.9%) individuals who identified themselves as “other” gender. The average age of the respondents was 36 (SD 10.3) years, with a range of 19 to 65 years and mostly between 19 and 39 years ($n=117$, 73.6%). More than half ($n=94$, 59.1%) of the participants disclosed their self-identification with AUD as “Very much like me” and the rest ($n=65$, 40.9%) disclosed it as “Somewhat like me.” In terms of prior awareness of AUDRA, 94 participants answered “Very much like me” and 65 participants mentioned “Somewhat like me;” prior awareness of AUDRA was exhibited by 136 (85.5%) participants.

Table 1. Sociodemographic characteristics of the respondents (N=159).

Variable	n (%)
Gender	
Male	111 (69.8)
Female	45 (28.3)
Other/undisclosed	3 (1.9)
Age (years)	
19-39	117 (73.6)
40-65	39 (24.5)
Undisclosed	3 (1.9)
Self-identification with AUD^a	
Very much like me	94 (59.1)
Somewhat like me	65 (40.9)
Prior awareness of AUDRA^b	
Yes	136 (85.5)
No	23 (14.5)
Total	159 (100)

^aAUD: alcohol use disorder.

^bAUDRA: alcohol use disorder recovery app.

Reliability and Validity of the Constructs

Table 2 describes the reliability and validity of the constructs. Internal validity was evaluated using the Cronbach alpha and CR, with the acceptable range falling between 0.6 and 0.7 [12]. The AVE values for all the constructs, except for facilitating conditions, were above 0.5, thereby indicating convergent validity. Note that the first item, FC1, pertaining to facilitating

conditions had to be removed because when FC1 was included along with the other items (FC2, FC3, and FC4), the CR value was very low (0.037). After removing FC1 from facilitating conditions, the CR value improved to 0.621. Therefore, 3 items related to facilitating conditions and all items pertaining to the other constructs were retained.

For the heterotrait-monotrait ratio, all comparisons were well under the recommended threshold of 0.85 and indicated satisfactory discriminant validity between the constructs (Table 2).

Then we compared the constructs to investigate differences between respondents. We compared respondents who identified “somewhat like me” and “very much like me” with AUD, as

shown in Table 3. The only significant difference was that the “very much like me” group found it slightly easier to use AUDRA.

Third, we compared respondents based on their prior awareness of AUDRA (Table 4). Respondents aware of AUDRA scored significantly higher on performance expectancy, effort expectancy, and social influence than respondents who had not.

Table 2. Construct reliability.

Construct	Cronbach alpha	Average variance extracted	Composite reliability
Performance expectancy (PE); loading	.678	0.593	0.812
PE1; 0.901			
PE2; 0.714			
PE3; 0.676			
Effort expectancy (EE); loading	.685	0.512	0.806
EE1; 0.794			
EE2; 0.650			
EE3; 0.759			
EE4; 0.648			
Social influence (SI); loading	.766	0.585	0.849
SI1; 0.720			
SI2; 0.749			
SI3; 0.764			
SI4; 0.824			
Facilitating conditions (FCs); loading	.412	0.407	0.621
FC2; 0.395			
FC3; 0.976			
FC4; 0.335			

Table 3. Level of identification with alcohol use disorder and participants’ mean scores on Unified Theory of Acceptance and Use of Technology constructs (N=159).

UTAUT ^a constructs	Self-identification with AUD ^b			
	Average out of 5	Very much like me (n=94)	Somewhat like me (n=65)	<i>P</i> value
Performance expectancy (3 items)	3.9	4.0		.56
Effort expectancy (4 items)	4.1	3.9		.03 ^c
Social influence (4 items)	3.7	3.8		.5
Facilitating conditions (4 items)	4.1	4.1		.8
Behavioral intention (1 item)	3.5	3.2		.57
Use behavior (1 item)	2.6	2.6		.93

^aUTAUT: Unified Theory of Acceptance and Use of Technology.

^bAUD: alcohol use disorder.

^cThe italicized *P* value is statistically significant.

Table 4. Prior awareness of the existence of smartphone alcohol use disorder recovery apps and participants' mean scores on Unified Theory of Acceptance and Use of Technology constructs (N=159).

UTAUT ^a constructs	Prior awareness of smartphone AUDRA ^b		
	Yes (n=136)	No (n=23)	P value
Average out of 5			
Performance expectancy (3 items)	4.0	3.6	.02 ^c
Effort expectancy (4 items)	4.1	3.7	.04
Social influence (4 items)	3.9	3.1	<.001
Facilitating conditions (4 items)	4.1	4.0	.45
Behavioral intention (1 item)	2.8	2.7	.74
Use behavior (1 item)	2.7	2.6	.75

^aUTAUT: Unified Theory of Acceptance and Use of Technology.

^bAUDRA: alcohol use disorder recovery apps.

^cThe italicized P value is statistically significant.

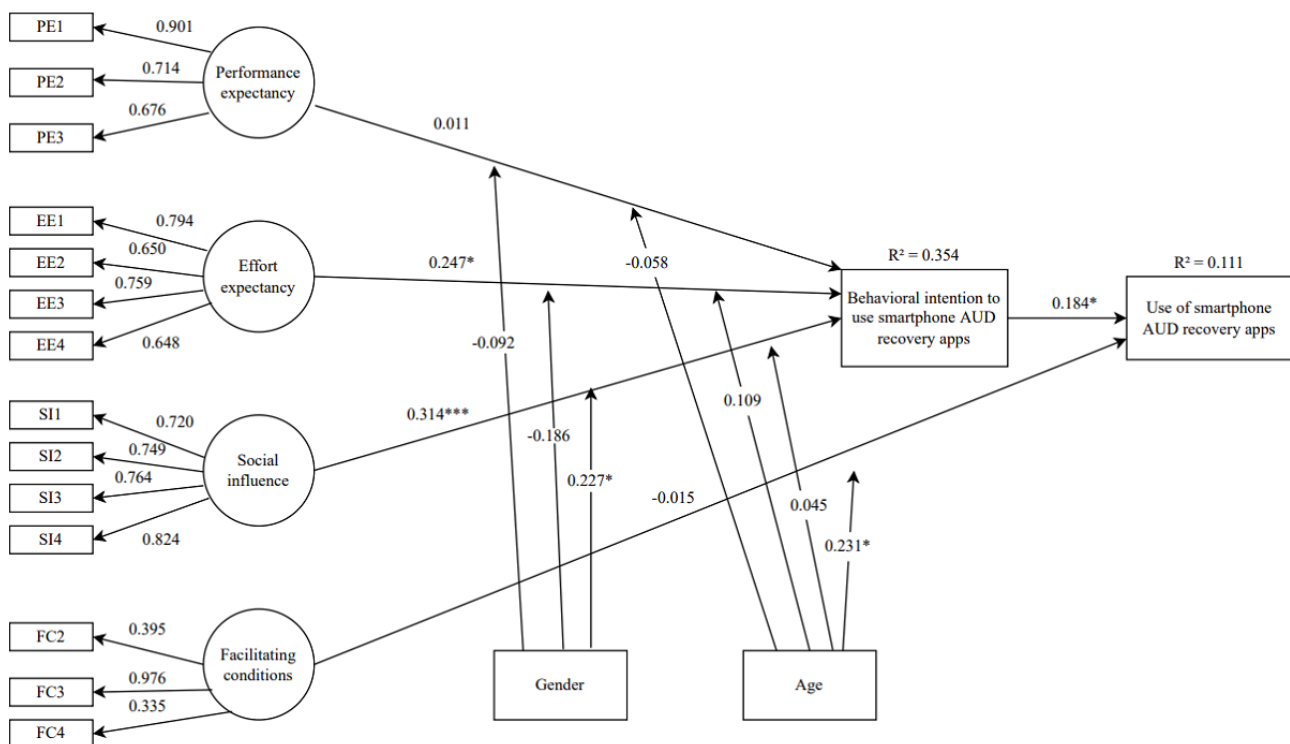
Structural Model to Identify the Behavioral Factors

To analyze the model fit, PLS-SEM was used. Figure 2 shows the path coefficients and the statistical significance of the relationships along with the coefficient of determination or the R² value.

Effort expectancy and social influence were significant predictors of behavioral intention to use smartphone AUDRA,

which itself predicted use. However, performance expectancy had no effect on behavioral intention. Gender moderated the effect of social influence, meaning that the effect of social influence on behavioral intention was more significant in women than in men. Facilitating conditions had no significant effect on use except for older users who were more likely to be influenced by facilitating conditions. Overall, the model explains 35.4% of the variance in behavioral intention and 11.1% of the variance in use behavior.

Figure 2. Complete model showing path coefficients and R². Statistical significance of the relationships (path coefficients): *P<.05; **P<.01; ***P<.001. AUD: alcohol use disorder; PE: performance expectancy; EE: effort expectancy; SI: social influence; FC: facilitating conditions.



Open-Ended Questions Regarding AUDRA

Open-ended responses provided further insight into participants' attitudes to AUDRA. Response rates on the 3 questions were between 35% and 67%. Privacy and security concerns were the most frequently given reasons by participants for not wanting

to use AUDRA. One respondent stated that "Privacy would be the only issue regarding using an app to help in recovery," whereas another pointed out "the potential of data tracking and possibility of using my information for profit." Other frequently given responses pointed to how "confusing" or "complicated" apps could be. Respondents also expressed their skepticism

over the efficacy of such apps in helping them with AUD recovery and noted specific user-unfriendly features, such as too many reminders, notifications, or advertisements: “Pop-ups asking me to rate and/or buy a pro version. Unsolicited communications.” The participants were also dissuaded from potential AUDRA use if there were technical glitches, or “bugginess,” with the apps.

In terms of what would make them want to use AUDRA, respondents asked if these apps would help them with abstinence and prevent relapse. Users often mentioned how a tracking feature (“track my days [without alcohol] and money savings”) helped them. On the contrary, many other users complained about the lack of a tracking feature in the apps they were using. Respondents also frequently cited the ability of apps to connect them with others through social networking features and with local resources, such as if they could “find a meeting close by” and “...Access to events happening through local AA chapter,” as major reasons why they would be encouraged to use the app.

Discussion

Principal Results

This study investigated the key antecedents of behavioral intention to use AUDRA among people with AUD. Generally, most of the 159 participants (n=136, 85.5%) were aware of AUDRA. This study confirms the role of effort expectancy and social influence as significant predictors of the intention to use AUDRA, similar to the findings of previous UTAUT studies on mHealth [6,14]. This was confirmed by open-ended answers suggesting that some of the main hurdles to use are technical glitches. However, performance expectancy was not found to significantly predict the intention to use from the final model. This is intriguing because this factor is considered the key determinant of technology usage in general [15-18]. However, it does not appear to apply to mHealth apps [14,19-21]. Other studies have highlighted that despite playing a major role, performance expectancy may not prove salient for mHealth apps when compared to other forms of technology and that effort expectancy plays a much more important role [8].

Facilitating conditions had no direct effect on use, but they were moderated by age. This suggests that facilitating conditions play a more important role as participants age. Other studies conducted with people aged over 60 [20] and 65 years [20,22] have also found a significant influence of facilitating conditions on the use of mHealth apps. Considering that our sample only had 1 participant aged over 60 years, this suggests that the importance of facilitating conditions may start at a younger age.

Participants' responses to the open-ended questions offer some insights into understanding these results. A major reason given by participants as to why they would not want to use AUDRA was that their privacy, confidentiality, or both could be compromised in any way. This fear has been echoed in many other studies in which respondents cited data privacy concerns as reasons for not using mHealth apps [23-27]. These concerns may have trumped other factors and dampened their intention to use these apps.

Future research should further investigate the factors leading to adoption of mHealth apps, such as concerns regarding privacy. This study also has implications for practitioners. With increased efforts being made to promote the use of AUDRA, designers should first focus on making their apps convenient and easy to use. For app designers, health care professionals, and health care authorities eager to promote the adoption of AUDRA, this study suggests focusing on social influence, ensuring that the use of AUDRA is supported and encouraged by the people who matter to potential users, including their family and general practitioners along with highlighting the positive experiences of other users in their network.

Limitations

This study has some limitations to be considered when interpreting the findings. We did not have enough respondents in the follow-up survey to measure use longitudinally. In addition, respondents self-identified their AUD status, and we could not verify it; however, previous studies, through test-retest validation, have suggested overall reliability with respect to such self-identification [28] associated with AUD. Many responses also had to be discarded. The gift certificate and the ease of access associated with the internet-based survey on the AUD Facebook groups may have attracted participants who were willing to break the survey rules and may explain the high number of surveys that had to be discarded. Finally, the sample size was relatively small, which comes with associated limitations, notably in terms of statistical power.

Conclusions

This study found that performance expectancy was not significant in explaining behavioral intention to use AUDRA. Instead, social influence and effort expectancy seem to be the key factors influencing the use of such apps. As apps extend their influence into highly intimate areas of our lives, the beliefs that determine the use of technology may be shifting away from utilitarian factors such as performance. Researchers and app developers alike should keep this in mind and consider the user environment and possibly privacy concerns when developing apps.

Acknowledgments

The researchers are grateful to the kind people at the addiction treatment centers from where data were partially collected for this study. This research would not have been possible without the feedback, generosity, and participation of the directors, staff, and the clients at these centers. The researchers also extend their thanks to the remaining participants who participated on the internet and contributed a significant portion of the data used in this study.

Authors' Contributions

RM collected the data and wrote the draft. JM provided directions and developed the overall research design. PN and HB reviewed the paper.

Conflicts of Interest

None declared.

References

1. Song T, Qian S, Yu P. Mobile health interventions for self-control of unhealthy alcohol use: systematic review. *JMIR Mhealth Uhealth* 2019 Jan;7(1):e10899 [FREE Full text] [doi: [10.2196/10899](https://doi.org/10.2196/10899)] [Medline: [30694200](https://pubmed.ncbi.nlm.nih.gov/30694200/)]
2. Kranzler HR, Soyka M. Diagnosis and pharmacotherapy of alcohol use disorder: a review. *JAMA* 2018 Aug;320(8):815-824 [FREE Full text] [doi: [10.1001/jama.2018.11406](https://doi.org/10.1001/jama.2018.11406)] [Medline: [30167705](https://pubmed.ncbi.nlm.nih.gov/30167705/)]
3. Fowler LA, Holt SL, Joshi D. Mobile technology-based interventions for adult users of alcohol: a systematic review of the literature. *Addict Behav* 2016 Nov;62:25-34. [doi: [10.1016/j.addbeh.2016.06.008](https://doi.org/10.1016/j.addbeh.2016.06.008)] [Medline: [27310031](https://pubmed.ncbi.nlm.nih.gov/27310031/)]
4. Becker S, Miron-Shatz T, Schumacher N, Krocza J, Diamantidis C, Albrecht U. mHealth 2.0: experiences, possibilities, and perspectives. *JMIR Mhealth Uhealth* 2014 Apr;2(2):e24 [FREE Full text] [doi: [10.2196/mhealth.3328](https://doi.org/10.2196/mhealth.3328)] [Medline: [25099752](https://pubmed.ncbi.nlm.nih.gov/25099752/)]
5. Vo V, Auroy L, Sarradon-Eck A. Patients' perceptions of mHealth apps: meta-ethnographic review of qualitative studies. *JMIR Mhealth Uhealth* 2019 Jul;7(7):e13817 [FREE Full text] [doi: [10.2196/13817](https://doi.org/10.2196/13817)] [Medline: [31293246](https://pubmed.ncbi.nlm.nih.gov/31293246/)]
6. Venkatesh V, Morris MG, Davis GB, Davis FD. User acceptance of information technology: toward a unified view. *MIS Quarterly* 2003 Sep;27(3):425. [doi: [10.2307/30036540](https://doi.org/10.2307/30036540)]
7. Venkatesh V, Thong JYL, Xu X. Consumer acceptance and use of information technology: extending the unified theory of acceptance and use of technology. *MIS Quarterly* 2012;36(1):157-178. [doi: [10.2307/41410412](https://doi.org/10.2307/41410412)]
8. Dwivedi YK, Shareef MA, Simintiras AC, Lal B, Weerakkody V. A generalised adoption model for services: a cross-country comparison of mobile health (m-health). *Gov Inf Q* 2016 Jan;33(1):174-187. [doi: [10.1016/j.giq.2015.06.003](https://doi.org/10.1016/j.giq.2015.06.003)]
9. Duarte P, Pinho JC. A mixed methods UTAUT2-based approach to assess mobile health adoption. *J Bus Res* 2019 Sep;102:140-150. [doi: [10.1016/j.jbusres.2019.05.022](https://doi.org/10.1016/j.jbusres.2019.05.022)]
10. Hair J, Hollingsworth CL, Randolph AB, Chong AYL. An updated and expanded assessment of PLS-SEM in information systems research. *IMDS* 2017 Apr;117(3):442-458. [doi: [10.1108/imds-04-2016-0130](https://doi.org/10.1108/imds-04-2016-0130)]
11. Hair JF, Sarstedt M, Ringle CM, Gudergan SP. *Advanced Issues in Partial Least Squares Structural Equation Modeling*. Thousand Oaks, California: SAGE Publications; Apr 2017.
12. Nunnally J, Bernstein I. *Psychometric Theory*. New York: McGraw-Hill; 1994.
13. O'Cathain A, Thomas KJ. "Any other comments?" Open questions on questionnaires – a bane or a bonus to research? *BMC Med Res Methodol* 2004 Nov;4:25 [FREE Full text] [doi: [10.1186/1471-2288-4-25](https://doi.org/10.1186/1471-2288-4-25)] [Medline: [15533249](https://pubmed.ncbi.nlm.nih.gov/15533249/)]
14. Sun Y, Wang N, Guo X, Peng Z. Understanding the acceptance of mobile health services: a comparison and integration of alternative models. *Journal of Electronic Commerce Research* 2013;14(2):183-200.
15. Apolinário-Hagen J, Hennemann S, Fritsche L, Drüge M, Breil B. Determinant factors of public acceptance of stress management apps: survey study. *JMIR Ment Health* 2019 Nov;6(11):e15373 [FREE Full text] [doi: [10.2196/15373](https://doi.org/10.2196/15373)] [Medline: [31697243](https://pubmed.ncbi.nlm.nih.gov/31697243/)]
16. Dwivedi YK, Rana NP, Jeyaraj A, Clement M, Williams MD. Re-examining the unified theory of acceptance and use of technology (UTAUT): towards a revised theoretical model. *Inf Syst Front* 2017 Jun;21(3):719-734. [doi: [10.1007/s10796-017-9774-y](https://doi.org/10.1007/s10796-017-9774-y)]
17. Taiwo AA, Downe AG. The theory of user acceptance and use of technology (UTAUT): a meta-analytic review of empirical findings. *J Theor Appl Inf Technol* 2013 Jan;49(1):48-58.
18. Williams MD, Rana NP, Dwivedi YK. The unified theory of acceptance and use of technology (UTAUT): a literature review. *J Enterp Inf Manag* 2015 Apr;28(3):443-488. [doi: [10.1108/JEIM-09-2014-0088](https://doi.org/10.1108/JEIM-09-2014-0088)]
19. Carlsson B. Internationalization of innovation systems: a survey of the literature. *Res Policy* 2006 Feb;35(1):56-67. [doi: [10.1016/j.respol.2005.08.003](https://doi.org/10.1016/j.respol.2005.08.003)]
20. Hoque R, Sorwar G. Understanding factors influencing the adoption of mHealth by the elderly: an extension of the UTAUT model. *Int J Med Inform* 2017 May;101:75-84. [doi: [10.1016/j.ijmedinf.2017.02.002](https://doi.org/10.1016/j.ijmedinf.2017.02.002)] [Medline: [28347450](https://pubmed.ncbi.nlm.nih.gov/28347450/)]
21. Nunes A, Limpo T, Castro SL. Acceptance of mobile health applications: examining key determinants and moderators. *Front Psychol* 2019 Dec;10:2791 [FREE Full text] [doi: [10.3389/fpsyg.2019.02791](https://doi.org/10.3389/fpsyg.2019.02791)] [Medline: [31920836](https://pubmed.ncbi.nlm.nih.gov/31920836/)]
22. Al-Gahtani SS, Hubona GS, Wang J. Information technology (IT) in Saudi Arabia: culture and the acceptance and use of IT. *Inf Manag* 2007 Dec;44(8):681-691. [doi: [10.1016/j.im.2007.09.002](https://doi.org/10.1016/j.im.2007.09.002)]
23. Atienza AA, Zarcadoolas C, Vaughon W, Hughes P, Patel V, Chou WYS, et al. Consumer attitudes and perceptions on mHealth privacy and security: findings from a mixed-methods study. *J Health Commun* 2015 Apr;20(6):673-679. [doi: [10.1080/10810730.2015.1018560](https://doi.org/10.1080/10810730.2015.1018560)] [Medline: [25868685](https://pubmed.ncbi.nlm.nih.gov/25868685/)]

24. Kao CK, Liebovitz DM. Consumer mobile health apps: current state, barriers, and future directions. *PM & R* 2017 May;9(5):S106-S115. [doi: [10.1016/j.pmrj.2017.02.018](https://doi.org/10.1016/j.pmrj.2017.02.018)] [Medline: [28527495](https://pubmed.ncbi.nlm.nih.gov/28527495/)]
25. Kotz D, Gunter CA, Kumar S, Weiner JP. Privacy and security in mobile health: a research agenda. *Computer* 2016 Jun;49(6):22-30. [doi: [10.1109/MC.2016.185](https://doi.org/10.1109/MC.2016.185)]
26. Krebs P, Duncan DT. Health app use among US mobile phone owners: a national survey. *JMIR Mhealth Uhealth* 2015 Nov;3(4):e101 [FREE Full text] [doi: [10.2196/mhealth.4924](https://doi.org/10.2196/mhealth.4924)] [Medline: [26537656](https://pubmed.ncbi.nlm.nih.gov/26537656/)]
27. Zhou L, Bao J, Watzlaf V, Parmanto B. Barriers to and facilitators of the use of mobile health apps from a security perspective: mixed-methods study. *JMIR Mhealth Uhealth* 2019 Apr;7(4):e11223 [FREE Full text] [doi: [10.2196/11223](https://doi.org/10.2196/11223)] [Medline: [30990458](https://pubmed.ncbi.nlm.nih.gov/30990458/)]
28. Selin KH. Test-retest reliability of the Alcohol Use Disorder Identification Test in a general population sample. *Alcohol Clin Exp Res* 2003 Sep;27(9):1428-1435. [doi: [10.1097/01.alc.0000085633.23230.4a](https://doi.org/10.1097/01.alc.0000085633.23230.4a)]

Abbreviations

AUD: alcohol use disorder
AUDRA: alcohol use disorder recovery apps
AVE: average variance extracted
CR: composite reliability
mHealth: mobile health
PLS-SEM: partial least squares-structural equation model
UTAUT: Unified Theory of Acceptance and Use of Technology

Edited by A Kushniruk; submitted 09.09.21; peer-reviewed by C Yu, B Chaudhry, V Buss; comments to author 07.11.21; revised version received 22.12.21; accepted 14.02.22; published 01.04.22.

Please cite as:

Menon R, Meyer J, Nippak P, Begum H

Smartphone Alcohol Use Disorder Recovery Apps: Cross-sectional Survey of Behavioral Intention to Use

JMIR Hum Factors 2022;9(2):e33493

URL: <https://humanfactors.jmir.org/2022/2/e33493>

doi: [10.2196/33493](https://doi.org/10.2196/33493)

PMID: [35363145](https://pubmed.ncbi.nlm.nih.gov/35363145/)

©Rijuta Menon, Julien Meyer, Pria Nippak, Housne Begum. Originally published in *JMIR Human Factors* (<https://humanfactors.jmir.org>), 01.04.2022. This is an open-access article distributed under the terms of the Creative Commons Attribution License (<https://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in *JMIR Human Factors*, is properly cited. The complete bibliographic information, a link to the original publication on <https://humanfactors.jmir.org>, as well as this copyright and license information must be included.

Original Paper

Evaluating User Feedback for an Artificial Intelligence–Enabled, Cognitive Behavioral Therapy–Based Mental Health App (Wysa): Qualitative Thematic Analysis

Tanya Malik¹, MA; Adrian Jacques Ambrose², MD, MPH; Chaitali Sinha¹, MA

¹Wysa Inc, Boston, MA, United States

²Vagelos College of Physicians and Surgeons, Columbia University, New York, NY, United States

Corresponding Author:

Tanya Malik, MA

Wysa Inc

131 Dartmouth St

Boston, MA

United States

Phone: 1 9874803442

Email: tanya@wysa.io

Abstract

Background: Digital mental health apps are rapidly becoming a common source of accessible support across the world, but their effectiveness is often influenced by limited helpfulness and engagement.

Objective: This study's primary objective was to analyze feedback content to understand users' experiences with engaging with a digital mental health app. As a secondary objective, an exploratory analysis captured the types of mental health app users.

Methods: This study utilized a user-led approach to understanding factors for engagement and helpfulness in digital mental health by analyzing feedback (n=7929) reported on Google Play Store about Wysa, a mental health app (1-year period). The analysis of keywords in the user feedback categorized and evaluated the reported user experience into the core domains of acceptability, usability, usefulness, and integration. The study also captured key deficits and strengths of the app and explored salient characteristics of the types of users who benefit from accessible digital mental health support.

Results: The analysis of user feedback found the app to be overwhelmingly positively reviewed (6700/7929, 84.50% 5-star rating). The themes of engaging exercises, interactive interface, and artificial intelligence (AI) conversational ability indicated the acceptability of the app, while the nonjudgmentality and ease of conversation highlighted its usability. The app's usefulness was portrayed by themes such as improvement in mental health, convenient access, and cognitive restructuring exercises. Themes of privacy and confidentiality underscored users' preference for the integrated aspects of the app. Further analysis revealed 4 predominant types of individuals who shared app feedback on the store.

Conclusions: Users reported therapeutic elements of a comfortable, safe, and supportive environment through using the digital mental health app. Digital mental health apps may expand mental health access to those unable to access traditional forms of mental health support and treatments.

(*JMIR Hum Factors* 2022;9(2):e35668) doi:[10.2196/35668](https://doi.org/10.2196/35668)

KEYWORDS

digital mental health; artificial intelligence; user reviews; cognitive behavioral therapy; CBT

Introduction

The World Health Organization estimates that 450 million people worldwide have a mental disorder and a mental health gap of 1:10,000 worldwide [1]. Another report identified financial constraints and lack of serviceability as structural barriers to treatment [2]. Despite considerable progress in access

to resources, the gap in mental health access, especially in industrialized countries, does not appear to have shifted [3,4]. Psychological and structural barriers to accessing mental health care, such as availability, convenience, stigma, and preference for self-care, persist and underscore the increased need for accessibility of mental health resources [5]. Digital mental health tools, such as apps and chatbots, allow for anonymity and convenience and can serve as important alternatives to bridge

the access gap [6]. The increasing availability and usability of mobile devices may create new opportunities for overcoming the existing barriers and limited access of traditional clinical service delivery and provide customized patient-centered interventions. Similarly, smartphones and other mobile technology may have the potential to reach a greater number of users and deliver reliable and effective services, regardless of location [7,8].

For bridging the mental health access gap, understanding user experiences and attitudes toward digital mental health apps is crucial. In the context of digital mental health, the Technology Acceptance Model posits that perceived ease of use and perceived usefulness of a given technology have a positive influence on user engagement, which is required for interventions to be effective [9]. For both patients and providers, Chan et al [10] proposed criteria to use in assessing mental health apps in 4 key domains: usefulness, usability, integration, and infrastructure. In addition, acceptability of a mobile app is defined as the perceived value, usefulness, and desirability [11]. As user engagement often can be suboptimal, users' attitudes toward the digital technology can reveal important insight into their engagement [9].

To further understand user engagement with artificial intelligence (AI)-guided digital mental health apps, this study aimed to understand user needs for impactful engagement with a digital mental health app (Wysa) by examining their user reviews. As a direct proxy for users' attitudes toward a digital mental health app, user reviews are typically voluntary, unsolicited, and openly available on a public forum, which may provide helpful evaluations and insights into the users' experiences and engagement. A previous qualitative analysis of user reviews on mental health apps identified design improvements, user expectations, unmet needs, and utility [12]. These user reviews are regarded as a comprehensive evaluation of the app from the user's own perspective, which provides rich insights into the app user experience [13,14]. In addition to understanding needs for engagement, this study planned to explore the perceived value, usability, and desirability of the app as a digital mental health tool [15,16].

Methods

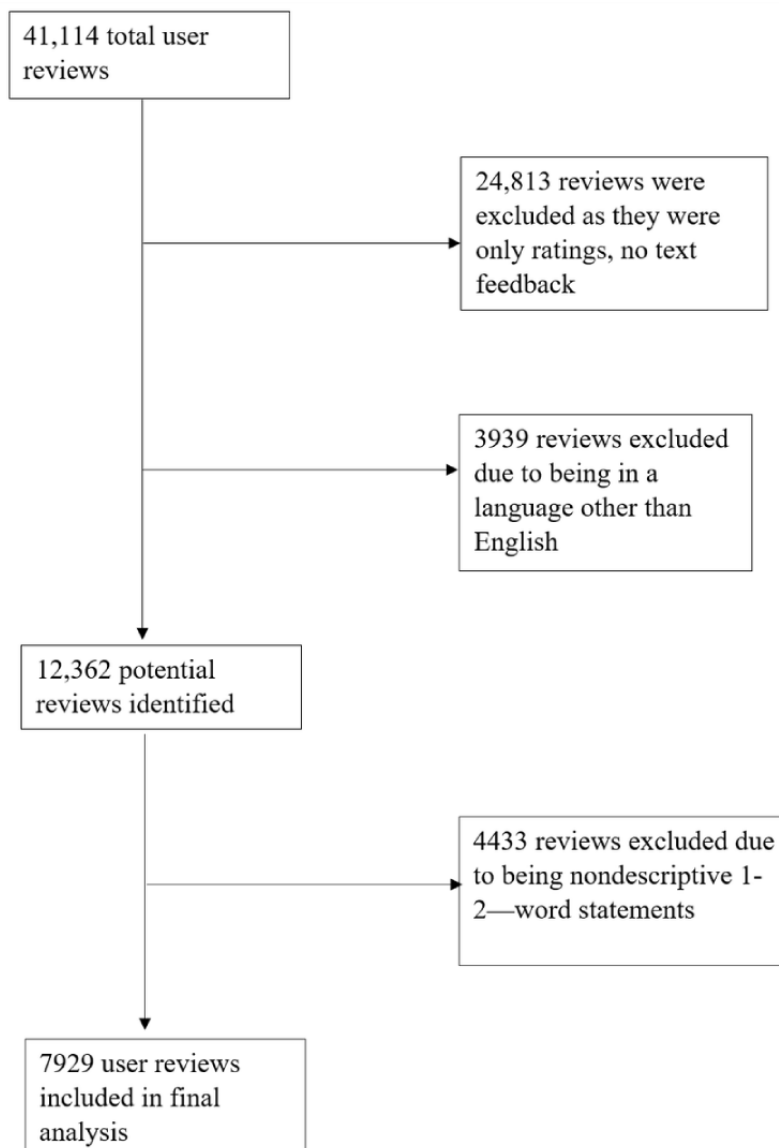
App Background

Wysa is an AI-enabled mental health app that leverages evidence-based cognitive-behavioral therapy (CBT) techniques through its conversational interface (chatbot). The app is designed by a team based out of India, the United Kingdom, and the United States. The app is designed to provide a therapeutic virtual space for user-led conversations through AI-guided listening and support, access to self-care tools and techniques (eg, CBT-based tools), as well as one-on-one human support. The app has demonstrated efficacy in building mental resilience and promoting mental well-being through a text-based conversational interface [17]. For the time period considered (1 year), the app received an overall 4.8/5 user rating on the Google Play Store and had been downloaded by more than 2 million people. The app also exists on the Apple App Store with a similar rating of 4.9/5 but with a smaller sample of qualitative reviews. Studies have shown Wysa as having the most evidence-based treatments among other smartphone apps [18], with conversations targeting specific problems and goals [19]. The app is anonymous [20] and safe [21] and rates highly on measures of app quality [22].

Study Design

For direct user feedback, the authors examined reviews posted on the Google Play Store between October 2020 and October 2021, during which time, 41,114 user reviews had been received. A duration of 1 year and the use of Google Play reviews were considered to ensure a sufficiently large sample. For the analysis of descriptive feedback (n=7929), the authors codified the reasons shared by the users for their rating. User feedback in languages other than English, blanks, as well as reviews that contained 1-2-word nondescriptive statements (eg, "Really nice!", "Awesome", "Not interested") were excluded (Figure 1).

The study's primary objective was to analyze feedback content to understand the users' experiences with engaging with a digital mental health app. As a secondary objective, the types of individuals providing feedback were also explored.

Figure 1. Diagram of the inclusion and exclusion criteria for user review analysis.

Analysis

Using a consolidated framework created by Chan et al [10], which was based on guidelines suggested by the Healthcare Information and Management Systems Society (HIMSS) and the US Federal Government for evaluating digital health apps, the written reviews were verbally grouped into the domains of the framework and further analyzed for specific themes within each [15]. To understand Wysa's capacity to currently help and engage users, the thematic analysis examined specific domains of (1) acceptability (eg, satisfaction, matching expectations of capabilities, likelihood to recommend, and level of interactiveness), (2) usability (eg, the ease, enjoyment, cultural, and demographic accessibility of use), (3) usefulness (eg, validity, reliability, effectiveness, and time required to obtain a benefit), and (4) integration (eg, security, privacy, data integration, and safety) [10,15].

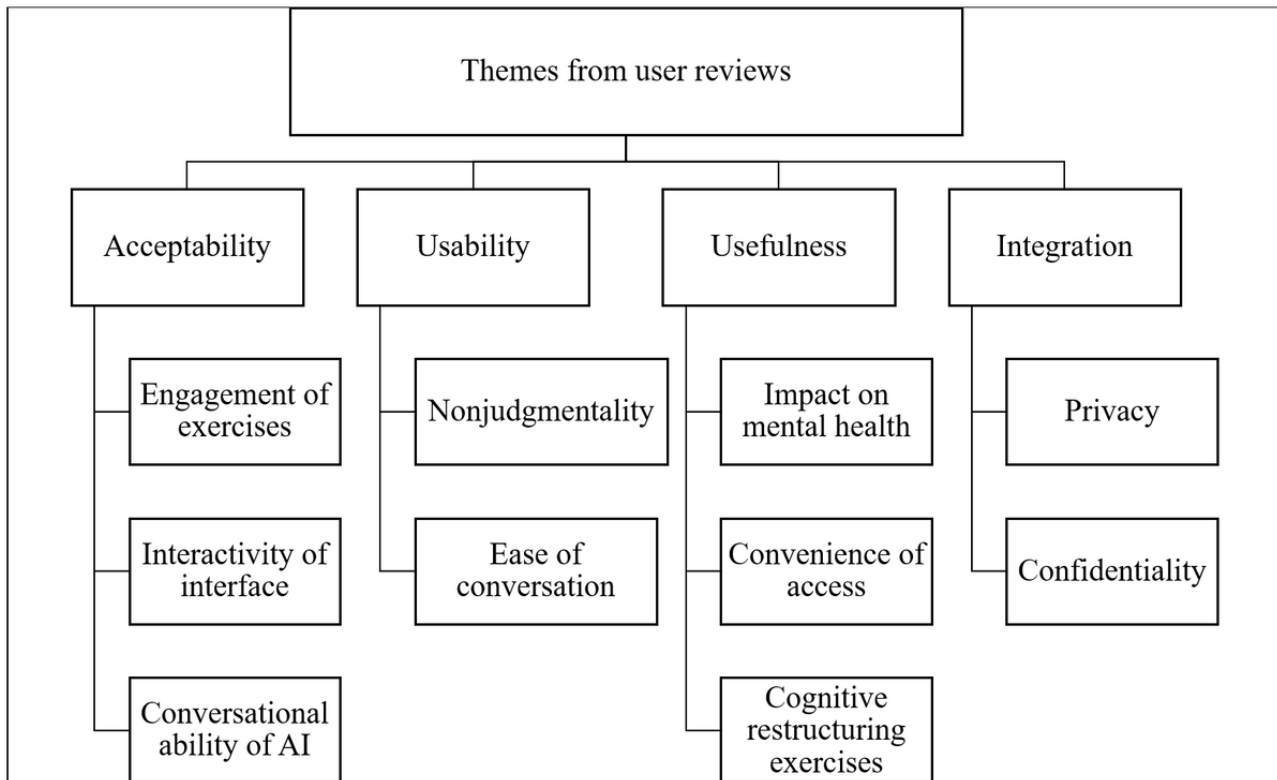
Each user review was evaluated and categorized into the nonmutually exclusive domains. The domain of acceptability included statements discussing likelihood to recommend the

app, frequency of use, impact of use, and reasons for use. For usefulness, mentions of what the app was being used for, specific uses (including tools and techniques), and time of use were included. Usability included mentions of ease of use, convenience, and interface features. Integration primarily consisted of reviews discussing data privacy, security, and anonymity.

The coding also enabled us to capture the emergence of the key characteristics of users who were able to receive mental health support due to increased accessibility.

Results

The reviews analyzed for this study were largely positive, with 6700 reviews (6700/7929, 84.50%) giving the app a 5-star rating and 2676 reviews (2676/7929, 33.75%) explicitly terming the app "helpful" or that it "helped." Of 7929 reviews, 251 (3.17%) had a less than 3-star rating and were termed as negative reviews. The themes under the evaluation criteria aimed to capture the user experiences (Figure 2).

Figure 2. Key themes from the reviews analyzed within the study. AI: artificial intelligence.

Acceptability

The acceptability of the app was identified through the themes emerging around engagement of exercises, interactivity of the interface, and conversational ability of AI. The users who reviewed the app rated it positively on acceptability when they found it interactive and conversational. Users reported that receiving appropriate responses to user conversations in the tools and techniques was valuable. For instance, a user compared it with other options available for self-care: “The interactive experience helped more than the journaling exercises I’ve done in the past.” Several users reported the variety of exercise-guided meditations, venting spaces, positive thinking exercises, and cognitive restructuring as important in their engagement. Another user commented: “It has such great features such as journaling and helping with anxiety, stress and sleep problems.” Additionally, the user reviews described the exercises as “educational,” “calming,” “relaxing,” and “functional.”

Users said that though “...Initially it felt silly to talk to an AI but it’s extremely well made, tailored for therapy.” Per users, the “warm, friendly, and encouraging” AI helped them recreate an environment of confiding in a friend, without having to confront the intimidation of speaking with a real person. For instance, a user mentioned “It’s really nice and I feel like I’ve been heard when others won’t listen, even if I am only talking to an AI,” and another user said it “made me feel loved and heard during a crisis.” Users also reported finding talking to the AI to be a “fun” experience, perhaps brought out by elements

that keep it light and accessible by including jokes, games, bitmap images (ie, GIFs), and other interactive agents.

Users reported the interactiveness of the app as central in keeping them engaged: “The app made me laugh with its silly jokes and play.” They also found the “easy” and “instinctive” interface as a central element in a positive experience of using the app to be “easy” and “instinctive” (Figure 3). Users found it comfortable to use Wysa for numerous aspects of their well-being (Figure 4).

They also mentioned being likely to recommend Wysa to others to help with sleep, managing stress, working through anxiety, as well as to “just talk to someone.” One user said, “it listens to you and helps relieve stress and also has a lot of coping mechanisms. I definitely recommend.” Some users discussed being able to share and rely on something for “regular” support, which further contributed to the acceptability of the app. A user exemplified this by stating:

Different people may find this app useful in different ways and it doesn't pressure you to do stuff if you aren't ready for it (no energy or not the right type). It's great even just as a sounding board, a place to organize your thoughts or make a to-do list, or a bit of a tiny friend in your pocket that's not judgmental and won't be tired of you.

However, some users did not find it helpful for their specific concerns and suggested further expansion to include these specific requirements. For instance, a user said, “Interesting concept, but it needs to learn to deal with more illnesses.”

Figure 3. Example of interactive interface.

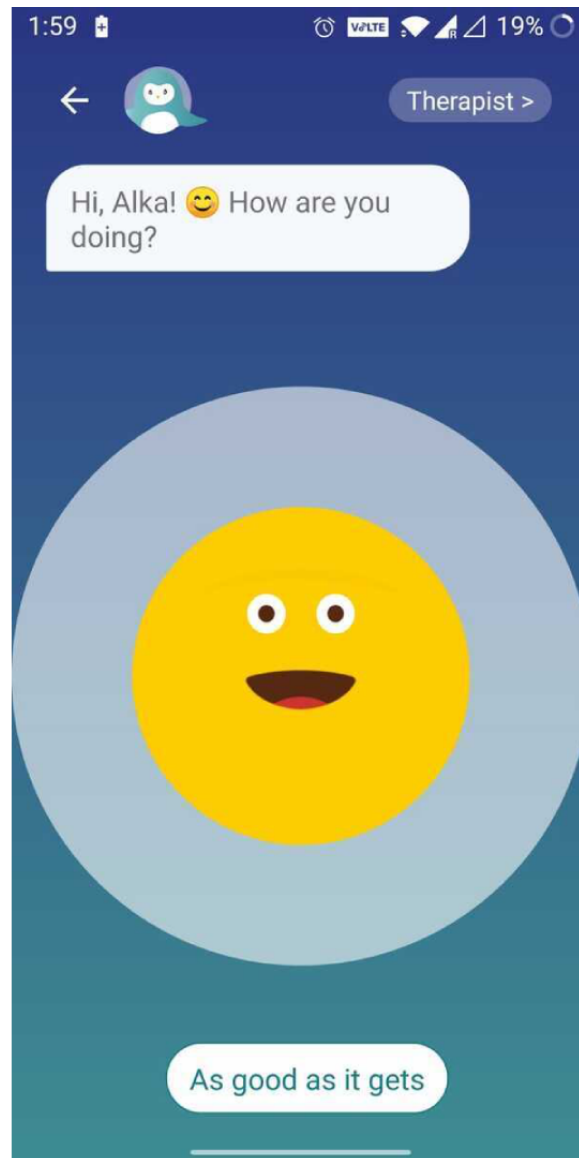
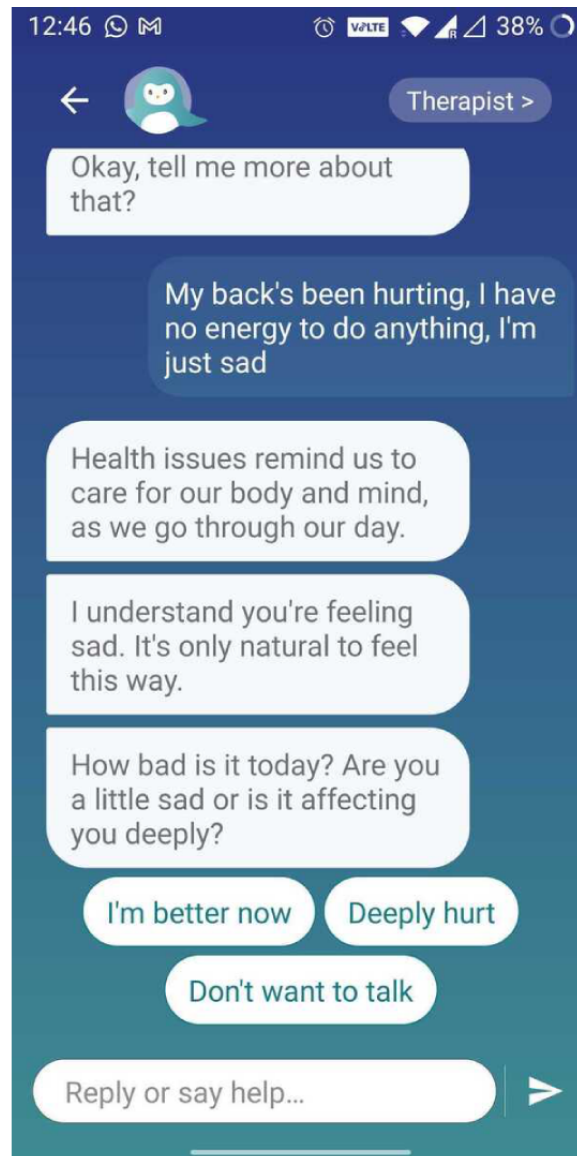


Figure 4. Example of empathetic conversation.

Usability

The usability of the app was presented through emerging themes of nonjudgmentality and ease of conversation. This domain was rated positively by users as they found it to be a “safe” and “nonjudgmental” environment that is easily accessible. This feature of the app was identified from user reviews such as:

It's nice to talk to someone completely objectively. Even in therapy you feel guilty if all you do is go on and on, as is human conditioning, but being able to talk it out with Wysa is great. No judgment. Don't have to feel weird about anything.

Reviews indicated that, by conversing with AI, the pressure of performance in front of a therapist was removed, which may allow a user to express themselves more freely. Users commented on the AI interface of the “cute and approachable penguin” as helpful in cultivating a nonthreatening environment: “I love it, it's just amazing, knowing that I can talk about my inner problems to a penguin without judgment ... I love that.” In fact, 201 users commented on the “no-judgment space” as a core component in making them feel safe and comfortable.

The app usage experience was also described as “...It feels like I'm talking to a real person ... Such a friendly interface.” Users appeared to be willing to adapt their expectation in order to continue benefiting from the app, with one user saying, “a little clunky at first but once you learn how to manage it it's very helpful.”

The most common negative review of the app was for repetition and a lack of comprehension by the chatbot, which made some users feel misunderstood and sometimes want to leave the app. Language limitations felt like a barrier to others who wanted to be understood more. They expressed a want for the app in native languages, including Italian, Spanish, French, and others, with one user saying, “The application is great, but it lacks the addition of other languages ... in order to facilitate its use by all layers of society.”

Usefulness

The app's usefulness was portrayed by themes such as impact on mental health, convenience of access, and cognitive restructuring exercises. User feedback discussed that the app provided a safe and open space to challenge one's thoughts and

feelings. The usefulness of the app in this regard is captured by its efficacy in dealing with mental health concerns. A user described their experience:

I have been struggling with depression since I was a child, and was terrified of reaching out for help. Finally a few weeks ago I hit rock bottom worse than ever before. I was really scared for a while. I was seeking some form of comfort or communication but didn't want to go to anyone, not to mention money is tight. This app really helped me when I needed it most. Who knew an AI penguin would cause me to sing again?

Providing a “safe” and “anonymous” place to process one's thoughts and emotions was identified by 107 users as highly impactful.

In specific clinical utility, users reported positive effects for anxiety (n=805), stress reduction (n=480), and depressed mood (n=400). In addition, 324 users reported app usage for posttraumatic stress disorder (PTSD) symptoms, fear, and sleep issues. Users identified numerous techniques and spaces offered as being especially helpful, such as physical activity exercises, sleep stories, meditations, cognitive restructuring, and reframing exercises. Users also commented on the affordability of the app as a way to bridge mental health access: “This app really helped me especially since I don't have access to any other useful form of therapy.”

The app would seem least useful when the chatbot felt limiting or was unable to fully understand the user. Some users facing a difficult time with the app would state, “Sometimes it's frustrating that an AI can't understand you that well,” and when it couldn't understand the user's dilemma, then it felt “empty and generic.”

Integration

The integration of the app was illustrated through the themes of privacy and confidentiality.

The app did not ask users to register themselves in any way to use the app and thus did not ask for personal details, such as demographic data. The anonymous and confidential nature of the app was a key reason for positive ratings in integration. Many users reported being satisfied with the privacy practices and finding the app “easy to trust.”

I feel really good knowing that I can talk to something completely private. I was feeling really down and I was pleasantly surprised. It was so simple yet so effective. I most definitely recommend it to someone who wants privacy and a healthy listening ear.

Characteristics of Users

The thematic analysis captured the emergence of the types of users who provided reviews in the app on Google Play Store and are also a representation of users who access digital mental health support such as Wysa. They were grouped by salient aspects of their expressed needs and concerns.

We identified 4 key groups: (1) those who self-reported having clinical issues, (2) those who reported being unable or unwilling

to open up to a real person, (3) those who are financially conscious, (4) and those who are unable to access mental health professionals. Use of the app for support through self-reported diagnosis and symptoms of depression, anxiety, panic disorders, and PTSD was mentioned by 1856 individuals. They primarily used the CBT techniques and meditations on the app as a form of self-care. Another application of the app is for individuals who feel uncomfortable talking to people in their lives or who don't have a reliable system with which to share their thoughts. They reported finding the AI-driven app useful in reducing the guilt and burden of opening up to a real person. Users also found the free nature of the app to be beneficial to reduce the burden of financial anxiety when considering mental health support. Numerous users (n=594) also reported using the app at times when they would be unable to access therapists, including when experiencing higher symptoms of depression and anxiety late at night.

Discussion

Principal Findings

This study represents one of the largest studies in understanding users' perceptions of a digital mental health app. It looked at the acceptability, usability, usefulness, and integration of a digital mental health app, by analyzing publicly available user feedback and reviews. This approach is unique for several reasons—first, it uses user feedback that was unsolicited by the developers and promoters and is delivered in a public forum, reducing the social desirability bias that could interact in other researcher-administered evaluations. Second, the robust sample size allowed for a deeper dive of user experiences, which was previously unexplored in other studies. This approach helped to recognize the types of users of mental health apps, which helped to identify strengths and weaknesses of digital mental health tools and allowed us to better understand the gaps in services provided.

The most important findings resulting from this study are the factors that contribute to higher engagement and acceptability for a digital mental health app. Users most consistently listed the “active and available listening” element as the key to foster acceptability with the digital mental health experience. The app further cultivated the therapeutic elements via the use of an AI-based chatbot with a friendly penguin user interface. In addition, the perceived nonjudgmentality and friendliness of this interface resulted in high usability and ongoing engagement with the app.

Understanding the user experience is important to ensuring meaningful usage and clinical utility [17]. Users strongly valued the anonymity and confidentiality of the app, which are valuable strengths in any therapeutic relationship [23]. Therapeutic bonds are fostered through trust, acceptance, empathy, and genuineness and are important for their role in the effectiveness of an intervention [23] and, in a digital environment, are created by human dialogue through a conversation agent [24].

With users providing a large majority of positive reviews, the acceptability and effectiveness of Wysa as a digital mental health tool have been established [25]. Digital mental health apps can

provide important benefits, especially for supporting individuals with subclinical psychiatric symptoms [26]. The findings of this study highlighted how digital mental health apps can significantly improve the accessibility and affordability of mental health support. The characteristics of users identified helped outline those who may access and benefit from the presence of mental health apps; for example, individuals managing social anxiety symptoms of speaking face to face can find significant therapeutic value through an AI-enabled tool. In addition, mental health apps may serve as augmenting or transitioning tools during times when traditional mental health services are limited, such as after office hours, in rural settings, or in between appointments and referrals.

Limitations

Limitations to the study include the source of data, as the Apple App Store data were not considered and only the reviews on Google Play Store were addressed in this study. Further, user reviews are taken at a single point in time, and thus evidence of changes in feedback are unavailable for consideration. No demographic information was collected aside from reviews being in English. Clinical scores of users were not identified,

which would otherwise have contributed to more direct understanding of the experience with the app in clinical populations. The study is also limited by lack of knowledge on the duration of app use or the rate of attrition among users due to app issues or other reasons.

Conclusions

This study utilized a user-led approach to understanding factors for engagement and helpfulness in digital mental health. User feedback was analyzed on domains of acceptability, usability, usefulness, and integration, and we found the app to be overwhelmingly positively reviewed. A key facet that emerged is the comfort and safe environment created by the nonjudgmental digital mental health tool that provides users with clinical and subclinical support. Further analysis revealed 4 predominant types of individuals who appear to be engaging in digital mental health support and who are infrequent users of face-to-face mental health services. Digital mental health apps can provide a valuable service to those unable to access mental health support. Future directions for digital mental health include improvements within the technology to cater to varied users, increasing its capacity to contribute to clinical utility.

Conflicts of Interest

TM and CS are employees of Wysa and hold equity in Wysa Inc. AJA declares no conflicts of interest related to this study.

References

1. Depression and other common mental disorders: global health estimates. World Health Organization. 2017. URL: <https://apps.who.int/iris/handle/10665/254610> [accessed 2022-03-30]
2. Rowan K, McAlpine DD, Blewett LA. Access and cost barriers to mental health care, by insurance status, 1999-2010. *Health Aff (Millwood)* 2013 Oct;32(10):1723-1730 [FREE Full text] [doi: [10.1377/hlthaff.2013.0133](https://doi.org/10.1377/hlthaff.2013.0133)] [Medline: [24101061](https://pubmed.ncbi.nlm.nih.gov/24101061/)]
3. Demyttenaere K, Bruffaerts R, Posada-Villa J, Gasquet I, Kovess V, Lepine JP, WHO World Mental Health Survey Consortium. Prevalence, severity, and unmet need for treatment of mental disorders in the World Health Organization World Mental Health Surveys. *JAMA* 2004 Jun 02;291(21):2581-2590. [doi: [10.1001/jama.291.21.2581](https://doi.org/10.1001/jama.291.21.2581)] [Medline: [15173149](https://pubmed.ncbi.nlm.nih.gov/15173149/)]
4. WHO MiNDbank: More Inclusiveness Needed in Disability and Development. World Health Organization. URL: https://www.who.int/mental_health/mindbank/flyer_EN.pdf [accessed 2022-03-30]
5. Jorm AF, Patten SB, Brugha TS, Mojtabai R. Has increased provision of treatment reduced the prevalence of common mental disorders? Review of the evidence from four countries. *World Psychiatry* 2017 Mar 26;16(1):90-99 [FREE Full text] [doi: [10.1002/wps.20388](https://doi.org/10.1002/wps.20388)] [Medline: [28127925](https://pubmed.ncbi.nlm.nih.gov/28127925/)]
6. Mohr DC, Riper H, Schueller SM. A solution-focused research approach to achieve an implementable revolution in digital mental health. *JAMA Psychiatry* 2018 Feb 01;75(2):113-114. [doi: [10.1001/jamapsychiatry.2017.3838](https://doi.org/10.1001/jamapsychiatry.2017.3838)] [Medline: [29238805](https://pubmed.ncbi.nlm.nih.gov/29238805/)]
7. East ML, Havard BC. Mental health mobile apps: from infusion to diffusion in the mental health social system. *JMIR Ment Health* 2015 Mar 31;2(1):e10 [FREE Full text] [doi: [10.2196/mental.3954](https://doi.org/10.2196/mental.3954)] [Medline: [26543907](https://pubmed.ncbi.nlm.nih.gov/26543907/)]
8. Torous J, Luo J, Chan SR. Mental health apps: What to tell patients. *Current Psychiatry* 2018;17(3):21-25 [FREE Full text]
9. Davis FD. Perceived usefulness, perceived ease of use, and user acceptance of information technology. *MIS Quarterly* 1989 Sep;13(3):319. [doi: [10.2307/249008](https://doi.org/10.2307/249008)]
10. Chan S, Torous J, Hinton L, Yellowlees P. Towards a framework for evaluating mobile mental health apps. *Telemed J E Health* 2015 Dec;21(12):1038-1041. [doi: [10.1089/tmj.2015.0002](https://doi.org/10.1089/tmj.2015.0002)] [Medline: [26171663](https://pubmed.ncbi.nlm.nih.gov/26171663/)]
11. Improving the User Experience. US General Services Administration. URL: <https://www.usability.gov/> [accessed 2022-03-30]
12. Nicholas J, Fogarty AS, Boydell K, Christensen H. The reviews are in: a qualitative content analysis of consumer perspectives on apps for bipolar disorder. *J Med Internet Res* 2017 Apr 07;19(4):e105 [FREE Full text] [doi: [10.2196/jmir.7273](https://doi.org/10.2196/jmir.7273)] [Medline: [28389420](https://pubmed.ncbi.nlm.nih.gov/28389420/)]
13. Alqahtani F, Orji R. Insights from user reviews to improve mental health apps. *Health Informatics J* 2020 Sep 10;26(3):2042-2066 [FREE Full text] [doi: [10.1177/1460458219896492](https://doi.org/10.1177/1460458219896492)] [Medline: [31920160](https://pubmed.ncbi.nlm.nih.gov/31920160/)]

14. Pagano D, Maalej W. User feedback in the appstore: An empirical study. 1961 Presented at: 21st IEEE International Requirements Engineering Conference (RE); July 15-19, 2013; Rio de Janeiro, Brazil p. 24-34. [doi: [10.1109/RE.2013.6636712](https://doi.org/10.1109/RE.2013.6636712)]
15. Gowarty MA, Longacre MR, Vilardaga R, Kung NJ, Gaughan-Maher AE, Brunette MF. Usability and acceptability of two smartphone apps for smoking cessation among young adults with serious mental illness: mixed methods study. *JMIR Ment Health* 2021 Jul 07;8(7):e26873 [FREE Full text] [doi: [10.2196/26873](https://doi.org/10.2196/26873)] [Medline: [34255699](https://pubmed.ncbi.nlm.nih.gov/34255699/)]
16. Marshall WL, Serran GA. The role of the therapist in offender treatment. *Psychology, Crime & Law* 2006 Aug 22;10(3):309-320. [doi: [10.1080/10683160410001662799](https://doi.org/10.1080/10683160410001662799)]
17. Torous J, Andersson G, Bertagnoli A, Christensen H, Cuijpers P, Firth J, et al. Towards a consensus around standards for smartphone apps and digital mental health. *World Psychiatry* 2019 Mar;18(1):97-98 [FREE Full text] [doi: [10.1002/wps.20592](https://doi.org/10.1002/wps.20592)] [Medline: [30600619](https://pubmed.ncbi.nlm.nih.gov/30600619/)]
18. Wasil AR, Venturo-Conerly KE, Shingleton RM, Weisz JR. A review of popular smartphone apps for depression and anxiety: Assessing the inclusion of evidence-based content. *Behav Res Ther* 2019 Dec;123:103498. [doi: [10.1016/j.brat.2019.103498](https://doi.org/10.1016/j.brat.2019.103498)] [Medline: [31707224](https://pubmed.ncbi.nlm.nih.gov/31707224/)]
19. Wasil AR, Palermo EH, Lorenzo-Luaces L, DeRubeis RJ. Is there an app for that? A review of popular apps for depression, anxiety, and well-being. *Cognitive and Behavioral Practice* 2021 Oct;1 [FREE Full text] [doi: [10.1016/j.cbpra.2021.07.001](https://doi.org/10.1016/j.cbpra.2021.07.001)]
20. Kretzschmar K, Tyroll H, Pavarini G, Manzini A, Singh I, NeurOx Young People's Advisory Group. Can your phone be your therapist? Young people's ethical perspectives on the use of fully automated conversational agents (chatbots) in mental health support. *Biomed Inform Insights* 2019 Mar 05;11:1178222619829083 [FREE Full text] [doi: [10.1177/1178222619829083](https://doi.org/10.1177/1178222619829083)] [Medline: [30858710](https://pubmed.ncbi.nlm.nih.gov/30858710/)]
21. Meadows R, Hine C, Suddaby E. Conversational agents and the making of mental health recovery. *Digit Health* 2020 Nov 20;6:2055207620966170 [FREE Full text] [doi: [10.1177/2055207620966170](https://doi.org/10.1177/2055207620966170)] [Medline: [33282335](https://pubmed.ncbi.nlm.nih.gov/33282335/)]
22. Myers A, Chesebrough L, Hu R, Turchioe MR, Pathak J, Creber RM. Evaluating commercially available mobile apps for depression self-management. *AMIA Annu Symp Proc* 2020;2020:906-914 [FREE Full text] [Medline: [33936466](https://pubmed.ncbi.nlm.nih.gov/33936466/)]
23. Nienhuis JB, Owen J, Valentine JC, Winkeljohn Black S, Halford TC, Parazak SE, et al. Therapeutic alliance, empathy, and genuineness in individual adult psychotherapy: A meta-analytic review. *Psychother Res* 2018 Jul 07;28(4):593-605 [FREE Full text] [doi: [10.1080/10503307.2016.1204023](https://doi.org/10.1080/10503307.2016.1204023)] [Medline: [27389666](https://pubmed.ncbi.nlm.nih.gov/27389666/)]
24. Følstad A, Brandtzaeg PB. Users' experiences with chatbots: findings from a questionnaire study. *Qual User Exp* 2020 Apr 11;5(1):1 [FREE Full text] [doi: [10.1007/s41233-020-00033-2](https://doi.org/10.1007/s41233-020-00033-2)]
25. Inkster B, Sarda S, Subramanian V. An empathy-driven, conversational artificial intelligence agent (Wysa) for digital mental well-being: Real-world data evaluation mixed-methods study. *JMIR Mhealth Uhealth* 2018 Nov 23;6(11):e12106 [FREE Full text] [doi: [10.2196/12106](https://doi.org/10.2196/12106)] [Medline: [30470676](https://pubmed.ncbi.nlm.nih.gov/30470676/)]
26. Mohr DC, Burns MN, Schueller SM, Clarke G, Klinkman M. Behavioral intervention technologies: evidence review and recommendations for future research in mental health. *Gen Hosp Psychiatry* 2013;35(4):332-338 [FREE Full text] [doi: [10.1016/j.genhosppsy.2013.03.008](https://doi.org/10.1016/j.genhosppsy.2013.03.008)] [Medline: [23664503](https://pubmed.ncbi.nlm.nih.gov/23664503/)]

Abbreviations

AI: artificial intelligence

CBT: cognitive behavioral therapy

HIMSS: Healthcare Information and Management Systems Society

PTSD: posttraumatic stress disorder

Edited by A Kushniruk; submitted 13.12.21; peer-reviewed by A Hashim; comments to author 18.02.22; revised version received 24.02.22; accepted 06.03.22; published 12.04.22.

Please cite as:

Malik T, Ambrose AJ, Sinha C

Evaluating User Feedback for an Artificial Intelligence-Enabled, Cognitive Behavioral Therapy-Based Mental Health App (Wysa): Qualitative Thematic Analysis

JMIR Hum Factors 2022;9(2):e35668

URL: <https://humanfactors.jmir.org/2022/2/e35668>

doi: [10.2196/35668](https://doi.org/10.2196/35668)

PMID: [35249886](https://pubmed.ncbi.nlm.nih.gov/35249886/)

©Tanya Malik, Adrian Jacques Ambrose, Chaitali Sinha. Originally published in *JMIR Human Factors* (<https://humanfactors.jmir.org>), 12.04.2022. This is an open-access article distributed under the terms of the Creative Commons

Attribution License (<https://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Human Factors, is properly cited. The complete bibliographic information, a link to the original publication on <https://humanfactors.jmir.org>, as well as this copyright and license information must be included.

Original Paper

Designing Tangibles to Support Emotion Logging for Older Adults: Development and Usability Study

Daniel Gooch¹; Vikram Mehta¹; Avelie Stuart²; Dmitri Katz¹; Mohamed Bennisar¹; Mark Levine³; Arosha Bandara¹; Bashar Nuseibeh^{1,4}; Amel Bennaceur¹; Blaine Price¹

¹School of Computing and Communications, The Open University, Milton Keynes, United Kingdom

²Department of Psychology, University of Exeter, Exeter, United Kingdom

³Department of Psychology, Lancaster University, Lancaster, United Kingdom

⁴Lero, the Science Foundation Ireland Research Centre for Software, University of Limerick, Limerick, Ireland

Corresponding Author:

Daniel Gooch

School of Computing and Communications

The Open University

Walton Hall

Milton Keynes, MK7 6AA

United Kingdom

Phone: 44 1908858234

Email: Daniel.Gooch@open.ac.uk

Abstract

Background: The global population is aging, leading to shifts in health care needs. In addition to developing technology to support physical health, there is an increasing recognition of the need to consider how technology can support emotional health. This raises the question of how to design devices that older adults can interact with to log their emotions.

Objective: We designed and developed 2 novel tangible devices, inspired by existing paper-based scales of emotions. The findings from a field trial of these devices with older adults are reported.

Methods: Using interviews, field deployment, and fixed logging tasks, we assessed the developed devices.

Results: Our results demonstrate that the tangible devices provided data comparable with standardized psychological scales of emotion. The participants developed their own patterns of use around the devices, and their experience of using the devices uncovered a variety of design considerations. We discuss the difficulty of customizing devices for specific user needs while logging data comparable to psychological scales of emotion. We also highlight the value of reflecting on sparse emotional data.

Conclusions: Our work demonstrates the potential for tangible emotional logging devices. It also supports further research on whether such devices can support the emotional health of older adults by encouraging reflection of their emotional state.

(*JMIR Hum Factors* 2022;9(2):e34606) doi:[10.2196/34606](https://doi.org/10.2196/34606)

KEYWORDS

older adults; health; emotion; affect; well-being; tangible interaction; TUI

Introduction

Background Context

The United Nations predicts that the global population aged 60 years and older will increase from 962 million in 2017 to 2.1 billion in 2050 and 3.1 billion in 2100, making this the fastest growing age group [1]. These demographic changes will significantly impact how we think about supporting the health and well-being of the population. Older people can face long-term disabilities and chronic conditions as well as mental

health difficulties [2]. For example, Age UK has noted that the number of over-50s experiencing loneliness is set to reach 2 million by 2025/6. This compares to around 1.4 million in 2016/7—a 49% increase in 10 years. For the purposes of this work, the term “older adults” is used to refer to anyone over the age of 50 years based on the recommendations of Age UK (the main charity working with older adults in the United Kingdom).

This increase in the older population will drive an increase in the need for carers and the costs of health care [3]. This has led to significant amounts of research into how to enable people to

age in place; “the desire and tendency of older persons to stay in their current dwelling units for as long as possible” [4]. Compared to other forms of care, aging in place is more cost-effective and preferred by many older adults [5]. This is because it can enhance many quality of life factors (eg, identity, autonomy, belonging, privacy, independence, social connections) [6,7].

There have been promising developments in the design of technology to support the physical health of an aging population [8-10]. However, there is increasing recognition of the link between well-being and “successfully” aging, which makes it important to improve the psychological well-being of older adults [11]. This necessitates mechanisms for the detection or logging of the older adult’s emotional state to either ensure that the older adult is happy or provide appropriate support when in emotional turmoil [12-14].

Although a wide variety of digital technologies have been developed for the monitoring of emotions [15-23], there is little work that explores such interfaces specifically for older adults [24]. In a review of apps for successful aging, no apps for monitoring emotions were identified [25]. Given that older adults have distinct cognitive, physical, and technical skills, alongside distinct emotional needs, it is necessary to consider the design of a system for recording the emotional state of older adults at home [12,26].

Many researchers argue that tangible user interfaces (TUIs) are ideal for use in domestic settings by older adults owing to both their acceptability in domestic settings and the comparatively quick learning curve [27-29]. TUIs allow the user to provide input to a digital system by manipulating physical objects (eg, moving them around or stretching and squeezing them). Similarly, output from the TUI interaction could be shown to the user through the manipulation of a physical object. TUIs have also been found to increase engagement with logging emotions, suggesting that this form factor could promote ongoing use [17]. A broad review of the TUI literature for supporting social interactions among older adults highlights that most papers conclude that TUIs are highly usable for older adults [30].

In previous laboratory-based work, we have demonstrated that nonfunctional prototypes of tangible devices allow older adults to log emotions and collect data comparable to validated psychological scales of emotion [31]. We build on this work by developing 2 of these nonfunctional prototype designs into tangible devices that can digitally record the logged emotions. Our field study with adults aged 51-85 years demonstrates the validity of logged data against existing scales of emotion, showing that tangible devices can provide data comparable to standard psychological scales in a home setting. We explored our participants’ experience of using the devices over a 6-week period. This provided an understanding of how users can appropriate the use of the devices as well as how key design characteristics are viewed. Our results highlight the potential of in-home tangible devices for recording the emotions of older adults and for supporting their emotional health through encouraging reflection of their emotional state.

Background Literature

By exploring previous approaches to logging emotion, we can identify key design properties that should be embedded in the design of tangible devices for logging emotions. Through exploring the literature on self-report scales of emotion, interfaces of logging emotion, and TUIs for logging emotion, we identify key design decisions and reflect on them when outlining the development of our TUI devices in the section “Designing tangible devices for logging emotions.”

It is important from the outset to distinguish between emotion and mood. Although both refer to phenomenological states, they differ in 2 key dimensions [32,33]. The first is time; emotions tend to be short-lived, whereas moods are more enduring. The second difference is that emotions are object-driven (ie, they relate to a specific object or experience), while moods are more general. The concepts are related; a person’s mood biases the emotions they experience and a person’s emotions contribute to the mood they are in. Throughout this paper, the term “mood” is used only when it is the term used by other researchers in their work. The terms “emotion” and “affect” are used interchangeably as is common practice [33].

Across all fields interested in emotional experience, there are 3 main approaches to detecting and measuring how people feel: physiological, behavioral cues, and self-report. This research is focused on self-reported measures of emotion. Although self-report measures have shortcomings, they provide the user with a level of control over the disclosure of their emotional state. This is important for older adults in having an active role in their health care needs [34,35]. Self-reporting emotions also has other benefits. From a well-being perspective, there is a rich literature on the benefits to an individual of emotional reflection and recording, which is commonly used as a therapeutic technique [36]. Studies are starting to show how technologically-mediated reflection and recording can improve well-being [37] and promote behavior change [38]. From a methodological perspective, a recent review of ecological momentary assessment of mood highlights the importance of self-reporting due to ecological validity and agency [39].

Self-report Scales of Emotion

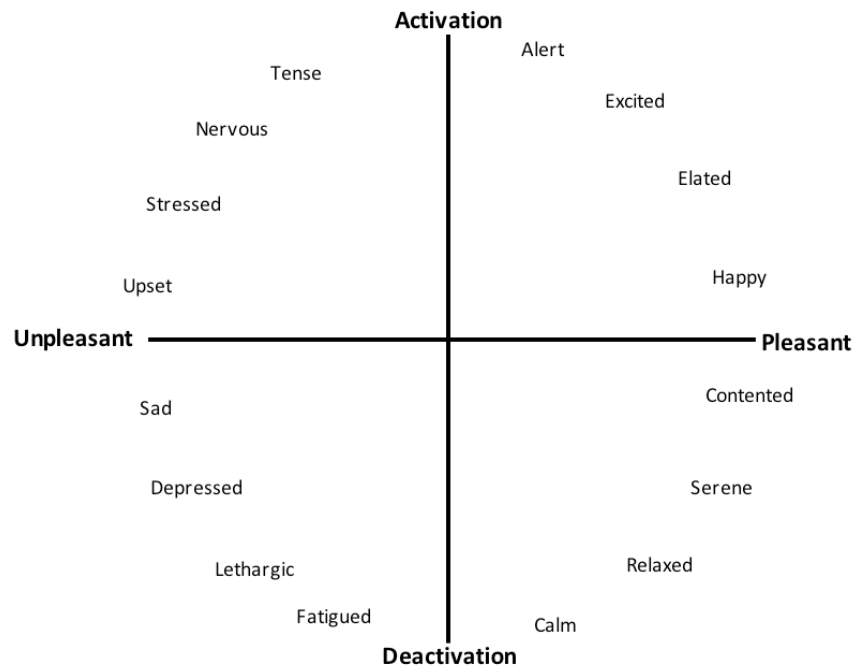
There are many different measures and scales focused on emotion in the psychology literature. Desmet et al [15] provide an excellent review of this literature. These measures predominantly coalesce around 2 concepts: valence (pleasure) and arousal (strength of feeling). Dominance is a third concept that is also sometimes used [40]. Proponents argue that these 3 dimensions can account for significant variances in people’s emotional experiences and collectively correspond to affect.

Russell’s 2D approach to conceptualizing emotion is one of the most popular measures of emotion [41,42]. He models emotion as a spatial distribution across 2 scales (valence and arousal) (see Figure 1). This approach argues that a spatial model provides a conceptual structure for related emotive concepts in such a way that allows the self-reporting of emotions [41]. A related approach uses emotive words to distinguish between related emotive states. One of the first commonly used robust

measures that took this approach was the Semantic Differential Scale, consisting of a set of 18 bipolar adjective pairs [43]. Each pair is then rated along a 9-point scale. Although heavily used,

the measure is extremely cumbersome to use, requiring 18 different measurement ratings for each stimulus. It also relies on an individual's English reading skills.

Figure 1. A schematic for the 2D structure of emotion from [41]. The valence scale runs left-to-right and the arousal scale runs top-to-bottom.



A variety of pictorial scales have also been developed. The Self-Assessment Manikin (SAM) is a classic example, made up of 3 pictorial scales: for affect (or valence), the pictures range “from a smiling, happy figure to a frowning, unhappy figure;” for arousal, the pictures range “from an excited, wide-eyed figure to a relaxed, sleepy figure;” and for dominance, the change is in the size of the figure, “a large figure indicates maximum control in the situation” [40]. Although SAM is straightforward to conceptualize, it is somewhat complicated to administer, particularly in terms of explaining the dominance dimension. Some have argued that the only intuitive scale is valence (positive vs negative facial expression) [15].

Alternative pictorial scales have aimed for even greater simplicity. For example, the “smileyometer” was developed as a single Likert-scale style set of emotive faces [44], while Desmet et al [15] generated 8 cartoon figures to represent key emotions. A questionnaire-based study with 191 participants suggests that their scale can provide robust and reliable assessments of individuals' emotions.

All of these scales were designed to be completed on paper. Given that we are designing an interactive technology for the collection of emotional data, we now explore the literature on interfaces for collecting self-reported emotions.

Interfaces for Logging Emotion

“A wide range of digital symptom monitoring tools exist, but there is a lack of evidence regarding their effectiveness in a health care context, particularly in the area of mental health” [45]. Much of the evidence that exists focuses on apps for people with mental health disorders (such as bipolar disorder) [46-51]. The findings from these studies highlight which design

properties have led to beneficial results and thereby provide insights into the design properties to utilize in the development of tangible alternatives.

An 8-week randomized trial of a suite of 13 mental health apps by Zhang et al [49] identified 3 distinct user behaviors: learning, goal setting, and self-tracking. Most importantly for our interests, participants who engaged in self-tracking experienced reduced depression symptoms. This is significant, as it suggests that logging affect can lead to improved emotional well-being. Zhang et al [49] also found that “greater amounts of engagement did not necessarily lead to greater reductions in depression.” This is an important design principle as it highlights that the device does not necessarily need to repeatedly harass users to enter data; as long as they engage with the system, they will receive some form of benefit.

True Colours is a digital tool for monitoring mood disorders. Used by over 36,000 individuals, it has formed part of 21 unique research and clinical service settings in the United Kingdom [45]. In addition to providing additional evidence of the efficacy of the digital logging of affect, the authors also note that the technology provides many advantages over hard copy symptom monitoring diaries, including the ability to prompt for input and the ability to easily visualize changes over time [45].

Chandrashekar [50] has reviewed meta-studies of the use of apps for people with depression, anxiety, and schizophrenia. In addition to demonstrating that these apps can have clinical benefits for these conditions, they also established some characteristics of high-efficacy apps. Among other features, providing a simple user interface and minimal usage reminders were highlighted as helping provide benefits to users.

Beyond these apps developed to help people with mental health disorders, there are a variety of interfaces that draw on self-report constructs of emotion to support the logging of emotion based on pictorial scales [15] or Russell's circumplex model [41,42]. None of these studies involved older adults, and the study focus was on exploring the developed design rather than the efficacy for users.

Fernández et al [52] developed a digital diary, specifically designed for older users. Users were encouraged to complete predefined questions about self-care and emotions answered on a tablet device. Fernández et al [52] focus on the usability elements of their design and field-tested the system with 10 participants aged over 60 years, who used the device for 5 days. Nine of the participants agreed that they would like to continue using the tool, and data collected from the study suggested that the simple act of logging was sufficient to prompt users to reflect about their day and how they were feeling.

Although the use of these interfaces has identified certain design properties as significant, they are not tangible devices. We now explore the sparse literature on TUIs for logging emotion to identify design properties specific to this interaction paradigm.

TUI for Logging Emotion

A small number of tangible interfaces has been developed to log emotions. The EmoBall [53] used an LED matrix grid to display "faces" with positive (smiling) or negative (frowning) faces. When the ball is pressed, the display shows a face depicting a different emotion; when the ball is pressed twice, the displayed emotion is logged and the ball vibrates. While evaluated through focus groups with 16 people, the study investigated the usability of EmoBall for people with "low digital competences" rather than its efficacy as a mood logging device.

In a different context, the subtle stone was developed to allow students to privately share their affect with their teacher within a classroom setting [54]. A ribbed rubber ball, the subtle stone contained 6 LEDs, which could display 7 separate colors. Each student could develop their own color/emotion mapping, and an emotion is selected by repeatedly squeezing the ball until the color is shown. This was field-trialed with 15 UK school students (aged 12-13 years) throughout 9 hours of German language lessons, with students reporting that the device "supported reflection on emotional experience by giving them a way of thinking about their emotions."

The Mood TUI was developed to make mood collection fun and engaging [17]. Designed as a cube with a different emoticon on each face, users select a mood by rotating the cube until the desired emoticon is facing upwards. Evaluated through discussion sessions with 32 participants, Sarzotti [17] concludes that there was interest in the design concept.

Jingar and Lindgren [55] took a design-oriented approach, co-designing TUIs to support the emotional health of older adults. Their interest was in how emotions could be communicated to a digital agent through tangible interactions. The variety of prototypes developed highlights the scope of the design space and the potential of TUIs to support older adults. Analyzing the data from their workshop, Jingar and Lindgren

[55] argue that the nature of TUIs means that they may be "intuitive and natural to use, and intrinsic motivation may be promoted" [55].

Our previous work has highlighted the value of TUIs, particularly for those older adults who have arthritis or other musculoskeletal difficulties. Arthritis is a common condition, particularly in later life [56], and musculoskeletal difficulties can limit an individual's ability to control a graphical user interface [57]. This makes tangible devices extremely suitable for use by older adults.

Research Objectives

Although there is substantial literature on developing apps, interfaces, scales, and measures for logging emotion, few are explicitly designed for older adults ([15-23] focus primarily on younger adults). We are specifically interested in designing tools to support older adults to log emotions; therefore, we draw on this work for inspiration. Given that research highlights the potential benefits of designing TUIs for older adults, we specifically focus on designing and developing novel tangible devices. Taking inspiration from existing paper-based scales of emotions, we explore what design properties are valued by older adults in the context of monitoring their emotional state. From the literature in the background section (see [Multimedia Appendix 1](#)) [16,17,20,31,45,49-55], the key design considerations that appear to have a significant impact on participants' use of the devices were to (1) minimize prompting, (2) ensure a clear mapping between the TUI interaction and the mood to be logged, (3) minimize fine grain movement, and (4) ensure that devices had a high-quality finish, suitable for use in a home location.

Designing Tangible Devices for Logging Emotions

We build on our earlier work on mood logging [31] to explore (1) whether digital TUIs can log emotional data comparable to validated psychological scales of emotion and (2) whether such devices would engage older adult participants and what their view of particular design characteristics were after using the devices in a home context. Thus, our first design decision was to focus on TUIs and convert the validated nonfunctional prototype designs into digital devices.

Key Design Decisions

Stepping back from the intricacies of particular device designs, it is necessary to discuss one of the underlying psychological practices that supports the efficacy of logging data: reflection. Reflection is a key part of all logging behavior. Manual data collection can support the process of reflection in action [58]. In the context of logging emotion, it is well-established that taking the time to consider your emotional state has benefits in itself, particularly in terms of someone deciding to change behavior based on their reflection [34,35,59-62].

Our second design decision was to provide the device users with no access to their recorded data during typical use. Users would only be shown their collected data at the end of the field deployment and if they asked to see it (to promote the transparency of the research). This stands in contrast to many self-logging devices but allows us to explore any benefits of

engagement with the data creation process, without confounding it with the benefits of reflecting on the historical data.

Our third design decision was to require minimal interaction [63,64], a design property that can help reduce the potential high burden of manual tracking. Given the perceived time burden of manual tracking [60], leading to high attrition rates [65], by minimizing the users’ interaction with the device, the potential time burden is also minimized.

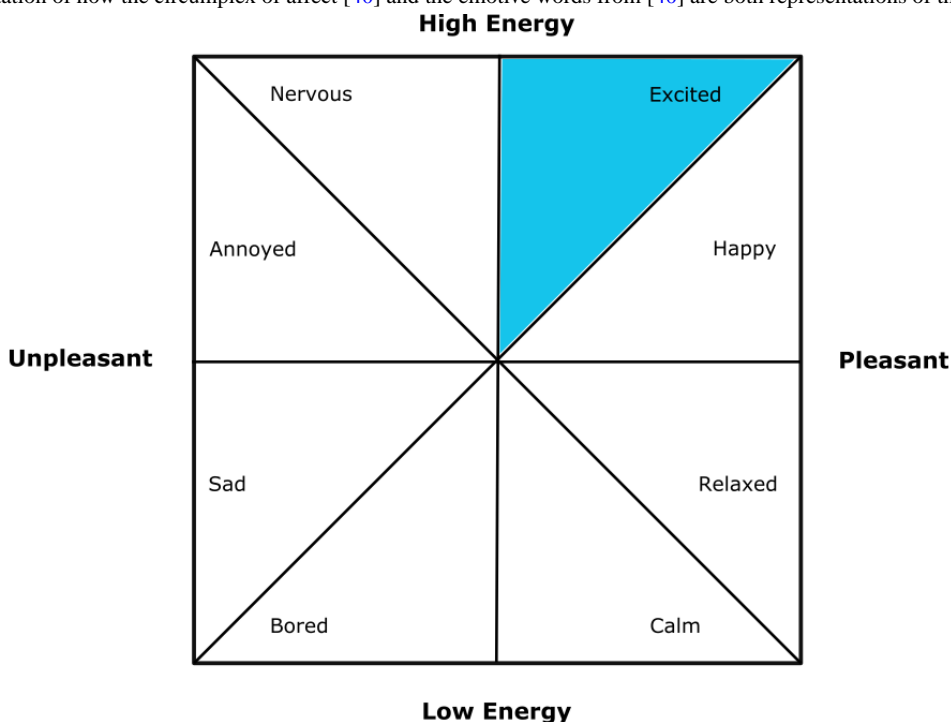
Selecting the Emotion Scale

The background section highlighted the wide range of available emotion scales. Our previous exploration of nonfunctional prototypes using 3 distinct scales indicated that 2 of the scales should be developed further into digital devices. The prototype

based on the emotive words from Russell’s circumplex were liked by users, given the simplicity of interaction and the speed of use. The prototype using the circumplex itself was liked by users, as it supported a more free-flowing process of reflection about their emotional state [31]. We decided to use these 2 scales of emotion.

Note that because these 2 scales represent the same conceptualization, analyzing the accuracy of logged data becomes easier. Figure 2 shows how the 2 scales can be considered to be somewhat equivalent. Taking the emotion of “excited,” the blue-highlighted octant can be taken to represent the emotion “excited” in the circumplex, and it is represented by the word “excited.”

Figure 2. A representation of how the circumplex of affect [40] and the emotive words from [40] are both representations of the same scale.



We chose to focus on developing devices that can record 8 emotions: happy, excited, nervous, annoyed, sad, bored, calm, and relaxed. These 8 emotions provide wide coverage over the range of potential emotions and are a commonly used subset of representative emotions [15].

Designing the Devices

Our previous work [31] focused on the development of nonfunctional prototypes of TUIs, which fulfilled the need to require minimal interactions [50,63,64]. As we have previously reported the design and development of these prototypes, here, we focus on the physical and electrical design of translating the nonfunctional paper prototypes into working digital TUIs. The resulting designs were named the Emotion Clock and the Emotion Board. These devices were developed by drawing on the design characteristics highlighted through the papers in the

background section, in constant conversation with experts at Age UK to ensure that the resulting designs would be appropriate for use by older adults.

Emotion Clock

The Emotion Clock arranges 8 emotive words around a clockface in accordance with Russell’s valence/arousal circumplex [41,42] (see Figure 3). A user selects an emotion by rotating the clock hand to the word describing the emotion they want to convey. The words are engraved into a wooden clock face, with the electronics hidden in a recess behind the clock face. The Emotion Clock has a diameter of 26 cm. Users were not instructed on how to use the hand. Although the clock allows users to record on a continuous scale, leaving the hand between 2 words, for the purposes of analysis, the nearest word to the hand position is recorded.

Figure 3. The Emotion Clock, using a subset of the emotive words in [40]. The emotion is set to Happy.



Emotion Board

The Emotion Board is a tangible representation of Russell's axes [41,42], using the color scheme from Rivera-Pelayo et al [20] (see Figure 4). The axes are labelled High Energy to Low Energy (top to bottom) and Feeling Bad to Feeling Good (left to right). A user moves a magnet around to select a position on the axes and thus represent an emotive state. Framed in wood,

there are 2 versions of the electronics behind the Emotion Board. The first version uses a custom piece of eTextiles, which is segmented to represent 16 sections of the axes (a high-arousal and low-arousal area for each of the 8 emotions). The second version uses an array of reed switches to achieve the same result but at a significantly lower cost. The board is approximately 26 square centimeters.

Figure 4. The Emotion Board, based on the Russell axes in [40] using the color scheme from [20]. The emotion is set to Calm.



Logging Behavior

The 2 devices adopted the same approach for how the underlying electronics capture the logged emotion. Both devices were controlled by a custom circuit board that could receive the logged mood from the device and transmit the log to a Raspberry Pi over Wi-Fi using the Message Queuing Telemetry Transport protocol. The Raspberry Pi was connected to the participants' router and could then send the log to our server using HTTPS.

The devices are powered using AA batteries rather than mains power. This allowed users to place the devices wherever they would like in their homes and improved the aesthetics of the devices by removing trailing wires. New batteries are sufficient to power the devices for at least 3 weeks. During the field trial, none of the participants' devices ran out of power. One implication of this decision is that the electronics must be low powered so that users do not have to repeatedly replace the batteries. As such, the electronics are programmed to capture the recorded data in a targeted way. Each log is recorded on a central server rather than locally on the device. This allowed us to monitor whether a deployed device was working. It also meant that we could keep an accurate record of the logged data without having to worry about the device being damaged and losing locally stored data.

Both devices "woke up" every 5 minutes to check the position of the clock hand or magnet. If the position had not changed (indicating no new emotion input), the device went back to sleep and nothing was recorded. If the position had changed, the device sent the new emotion to our servers over Wi-Fi and recorded it locally (replacing the previously recorded emotion). The device would try to send the data to the servers up to 10 times before returning to sleep; if it had not successfully sent the emotion, it would attempt to send the locally recorded emotion the next time it woke up—this would continue until the batteries ran out.

Following advice from the literature, the devices prompt users to log their emotions regularly but infrequently to ensure sufficient reflection without placing an undue burden on users [45,49,50]. The devices beeped at noon and 6 PM for 5 seconds irrespective of how many inputs were given by the user for that date. The devices did not beep at any time outside this window. To encourage at least 1 logging action per day, between noon and 6 PM, the device beeped on every hour until a mood was logged. In designing this protocol, the disruption of users was minimized while prompting them to think about their emotions.

Methods

Ethics Approval

Our study was designed in accordance with our University's code of ethics and approved by the Open University Human Research Ethics Committee (HREC/3343/Gooch).

Device

Each of the devices was piloted and was found to induce no discomfort. Participants had the right to refuse to use either of the devices, and it was possible for participants to immediately end their use of a device if they experienced any discomfort.

None of the participants opted to do so. We had 2 key concerns in exploring the value of the developed devices. The first is whether participants could accurately record their emotional state through the prototype. The second concern was to explore how our participants used the devices and their view on the design characteristic embodied within the devices.

Procedure

A field-trial approach was used to evaluate the devices over a period of 6 weeks. This involved each participant taking part in a pre-study session, a mid-study session, and an exit-study session. Each of these sessions took place at a participant's home and were audio recorded. The sessions lasted between 25 and 54 minutes (mean 28 minutes). Each session was one-to-one between a researcher and participant. Each participant used both devices for 3 weeks. The ordering of which device was used first was counterbalanced between participants as much as possible, although more Emotion Clocks had been manufactured, meaning the majority of participants ($n=7$) used this device first. The semistructured interview script can be found in [Multimedia Appendix 2](#). The procedure was as follows.

Pre-study Session

Sessions began by the researcher explaining that the purpose of the study was to explore new ways of logging emotion and highlighting that no personal emotional experiences would be logged. Informed consent was then collected. Subsequently, this session comprised the following activities: (1) initial data collection, (2) device orientation, (3) emotion logging calibration, and (4) device setup. Each of these activities is described below.

1. Initial data collection: Some basic demographic information was collected from the participant, as well as conducting a short interview regarding any existing logging behaviors (such as keeping a diary), their use of logging technology (such as a Fitbit), and what prompted the participant to take part in the trial.
2. Device orientation: Participants were given a brief explanation of one of the devices (counterbalanced between participants) and how they represent the 2 dimensions of emotion. The researcher answered any questions the participant had regarding the device.
3. Emotion logging calibration: The main element of the pre-study session was to gather data as to whether participants could log emotions using the selected device with the same accuracy as with the standardized paper-based scales. To ensure coverage across different emotional states, standardized emotive vignettes were used. The Affective Norms for English Text (ANET) vignettes are linked to known SAM scores, giving us a known emotion associated with each vignette [66] (referred to as the expected vignette emotion). These texts have previously been used in studies of emotional interfaces [16], as well as with our previous nonfunctional prototypes [31]. For each of the 8 emotions (happy, calm, nervous, excited, sad, relaxed, bored, and annoyed), a short vignette with SAM scores corresponding to that emotion was selected. A condition of using the ANET vignettes is to keep them confidential; so, we are unable to republish them. To illustrate the tone of the

vignettes, these 2 examples were written by the first author: (1) “You receive a letter informing you that you have won a holiday to the Caribbean in the quiz you entered last week” (excited) (2) “You discover that your best friend has been diagnosed with a serious illness” (sad). Participants were provided with the vignettes in a randomized order. Having read the text, participants were asked which emotion was portrayed by the vignette. This description is referred to as the participant description. For all of the vignettes, all of the participants provided a synonym of one of the 8 emotions (eg, thrilled becomes excited). The participant description allows us to test that the emotion logged by a participant through the prototype matches the emotion the participant wanted to log. Participants were then asked to record the emotion from the vignette through the prototype. The researcher recorded the result for the prototype alongside the time taken by the participant to record the emotion. Completing this exercise prior to setting the device up means that the logged emotions do not include this initial test.

4. Device setup: The prestudy session ended with the researcher setting the device up within the participants’ home for them to log their emotions for 3 weeks. Participants were instructed that they could place the device wherever they wanted within the home. In terms of use, participants were told that “the device will prompt you to input your emotions twice a day. You can provide more inputs if you wish to.”

At the end of the session, participants were provided with contact details and informed that they could contact us at any time if they were experiencing problems or wanted to talk about the study. We could remotely monitor whether the devices were working correctly by checking the server holding the logged emotions.

Midstudy Session

The focus of the midstudy session was to swap over the 2 devices at 3 weeks after the prestudy session. The session started with an audio-recorded wrap-up interview for the device the participant had been using for 3 weeks. The interview covered aspects such as exploring whether the participant had noticed an impact on how they felt, what their general thoughts about the device were, and specific questions regarding the prompting, the aesthetics, the difficulty of interaction, and whether they would hypothetically be willing to share the emotion data they had recorded. Having completed the interview, the researcher swapped over the devices and then repeated the prestudy session with the participant for the second device.

Exit-Study Session

Three weeks after the midstudy session, the exit-study session concluded the study and compared the experience of using the

2 devices. The session started with a wrap-up interview for the device the participant had been using for 3 weeks, following the same procedure as for the midstudy session. The session concluded by asking participants to complete a short interview, which was audio recorded. Participants were asked about their general thoughts about the idea of recording their emotions, how hard they found each prototype to use, how hard each prototype was to understand, and their opinions about having a similar device in their home. Further questions explored whether participants continued to be interested in logging how they felt; comparing the 2 devices in terms of use, aesthetics, and how hard they found each prototype to use; and any changes the participant could suggest for improving either of the devices. The study ended with a short debrief, during which time participants were thanked. Participants were shown graphs of their mood data for full disclosure of the collected data. Participants were provided with a £30 (US \$39) honorarium for taking part in the study.

Analysis

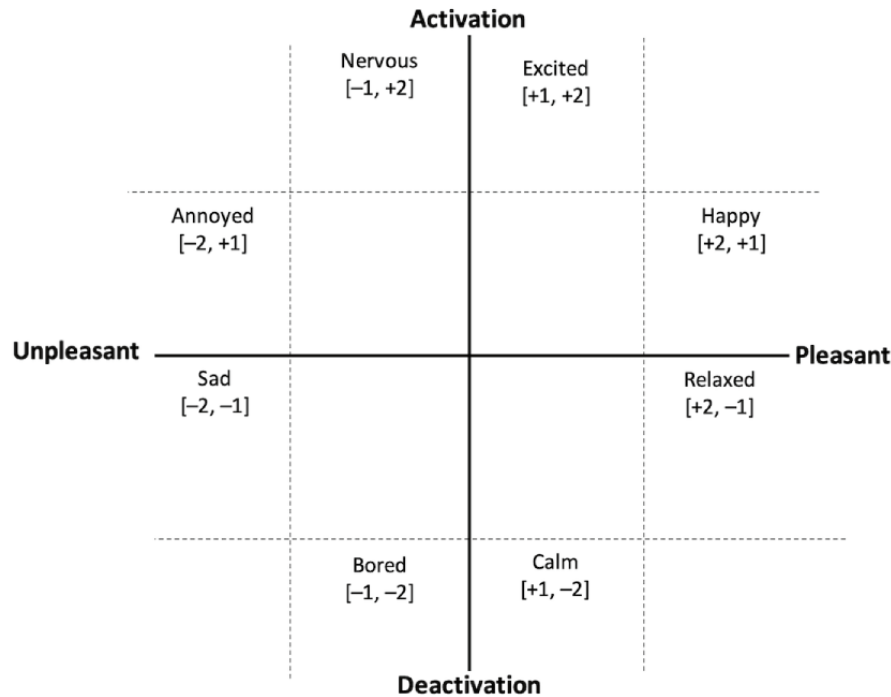
In analyzing the data from the study, we had 2 main questions. The first relates to the accuracy of the prototypes: could participants log the emotion they want to log through the prototype devices? The second was to explore our participants’ use of the devices and consider their response to the design characteristics embodied by the devices.

Accuracy of the Prototypes

The data from each of the prototypes can be analyzed categorically and ordinally, as outlined previously [31]. As categorical data, there is “ground truth” for each vignette because each vignette is taken from a validated set of emotive texts. Therefore, the emotion the vignette should be provoking in our participants is known (the *expected vignette emotion*). We also have the *participant description*, the emotion the participant believes each vignette expresses. To determine whether the prototypes allow participants to log the emotion they wanted to record, Cohen kappa is used to compare the emotion recorded through the prototype against (1) the *expected vignette emotion* and (2) the *participant description*. Cohen kappa ranges from no agreement ($\kappa=0$) to complete agreement ($\kappa=1$) [67].

A problem with treating the data as categorical is that it removes any connection between the different emotions. For example, if a participant records “happy” instead of “excited,” that is a closer match than if they record “sad.” An alternative way of conceptualizing the data is as 2 ordinal scales. Each of the prototypes uses a scale based on Russell’s circumplex of affect (see Figure 2); therefore, each emotion can be represented as a pair of figures ranging from -2 to $+2$ for both valence and arousal (see Figure 5).

Figure 5. A representation of how the emotions can be given ordinal values on the circumplex of affect.



As an indicator of similarity, it is possible to calculate the Euclidean distance by calculating the distance between 2 matrices (the expected emotional values and the actual emotional values), with each matrix being formed of the valence and arousal values. The distance reflects the size of dissimilarity between the expected emotions and the recorded emotions; the more dissimilar, the greater the distance between them. The

Euclidean distance between 2 observations is the length of the line between them. The equation in Figure 6 is used to calculate the distance across all samples. In both the categorical Cohen kappa and the ordinal Euclidean distance, we are not interested in the statistical performance *per se*. Instead, we are looking for confirmation that the prototypes allowed participants to log the emotion they wished to record.

Figure 6. The equation for calculating Euclidean distance.

$$D_i = \sqrt{\sum_{j=1}^n (A_j - B_j)^2}$$

Analyzing Participants' Views

The interviews were audio recorded and transcribed. An inductive open coding approach was used to identify concepts and themes within the interview transcripts [68]. The transcripts were subjected to a line-by-line analysis by the first author, who had not interviewed any of the participants. Through this initial analysis, concepts were identified and labelled within the data. No codes existed prior to the analysis; they were created through constant comparison of the data and the application of labels to the text.

These codes were subsequently categorized into unifying themes by the first author. These themes were there discussed in conjunction with the 3 authors who had interviewed the participants, to ensure that the developed themes corresponded with their interpretation of the data, informed by the interviews they had participated in.

Results

Recruitment

Eleven participants were recruited to take part in the study. The study was interrupted by the COVID-19 pandemic. This meant participant #9 could not fully complete the study, as it was not possible to switch the devices over and they only used the Emotion Clock. A further 2 participants (participant #10/participant #11) could not start the study owing to difficulties of setting the devices up within their homes. Two other participants were somewhat impacted by the pandemic, with participant #5 and participant #8 using their second device, as the United Kingdom went into lockdown. It is unknown whether this had an impact on their logging behavior. We have full data from 8 participants, and partial data from participant #9.

Participants had to be aged over 50 years, be fluent in English, and to have no significant cognitive impairments. Participants' ages ranged from 51 to 85 years (mean 69 [SD 11.9] years).

Seven of our 9 participants were females. All 9 participants had English as their first language. None of the participants reported a history of mental health concerns. Participants were recruited through Age UK Exeter (participant #6, participant #7, participant #9) or personal contact with the authors (participant #1-5, participant #8) through word of mouth or previous

participation in other studies. None of the participants had disruptive physical difficulties or cognitive impairments. [Table 1](#) shows the demographics of our participants.

We found no differences in our analysis between those participants who received the Emotion Clock first and those who received the Emotion Board first.

Table 1. Demographics of our participants.

Participants	Age (years)	Gender	First device
Participant #1	69	Female	Emotion Clock
Participant #2	74	Female	Emotion Board
Participant #3	69	Female	Emotion Clock
Participant #4	51	Male	Emotion Clock
Participant #5	54	Female	Emotion Clock
Participant #6	85	Female	Emotion Clock
Participant #7	60	Male	Emotion Board
Participant #8	79	Female	Emotion Clock
Participant #9	80	Female	Emotion Clock

Accuracy of the Logged Emotions

Using standard ANET vignettes provides baseline data of the emotion associated with the vignette, while the *participant description* states what emotion the participant wanted to record. Both can then be compared against the emotions recorded through the 2 prototypes.

The first stage of this comparison is to examine the results as categorical data. [Table 2](#) presents the results from calculating Cohen kappa for each prototype, comparing the emotion recorded in the prototype against (1) the expected result based on the ANET vignette scores and (2) the participant-described emotions. The results show at least moderate agreement (all kappa values > 0.5 at $P < .001$) [69], with the Emotion Clock demonstrating strong agreement.

Table 2. Cohen kappa values for each prototype.

Prototype	Expected vignette emotion	Participant description emotion
Emotion Clock	0.79	0.91
Emotion Board	0.5	0.5

Examining the results as ordinal data, we calculated the Euclidean distance between the valence/arousal values collected through the prototypes and the expected valence/arousal from the vignettes. The Euclidean distance between the values collected through the prototypes and the participant's description

of the vignette was also calculated. [Table 3](#) shows the Euclidean distances for each of the prototypes. To interpret these figures, it is important to note that there are 64 data points (8 vignettes from 8 participants) on 2 scales running from -2 to +2.

Table 3. The Euclidean distance for the valence and arousal data recorded through each interface compared against the expected data from the vignette and the participant description.

Prototype	Vignette total distance	Participant description total distance
Emotion Clock	21.65	18.35
Emotion Board	58.45	21.40

To contextualize the data, we also calculated what the Euclidean distance would be if, for a given interface, all participants were 1 emotion out (see [Figure 5](#), eg, the expected emotion was "excited" and the participant records "happy"). Such a scenario provides a Euclidean distance of 90.51. We also calculated what the Euclidean distance would be if, for a given interface, all participants provided the opposite emotion (eg, the expected emotion was "happy" and the participant records "sad"). Such a scenario provides a Euclidean distance of 286.22. Compared

against these contextual calculations, our results in [Table 3](#) show strong-to-moderate agreement between the expected emotion and the recorded emotion. This suggests that the disagreements between expected emotions and recorded emotions noted by the Cohen kappa results were not large discrepancies (eg, logging "happy" instead of "sad") but small (eg, logging "excited" instead of "happy").

Consistent with the kappa results, these results show a clear difference in the accuracy of the prototype responses with the

emotions logged through the Emotion Clock being the closest to both the vignette and *participant description* values.

Participant Use of the Devices

Usage Behaviors

Having established the accuracy of the devices, we considered the ways in which our participants used the prototypes. Our 9

participants recorded 1085 emotions across the 42-day study (see Table 4). The graph in Figure 7 shows the number of emotions recorded by each participant by study week. This shows some indication of novelty effects (with a high peak for most participants in week 1 and then, a general decline), but the number of emotions recorded is relatively consistent over time.

Figure 7. A graph showing the number of logs made by each participant by study week. P: participant.

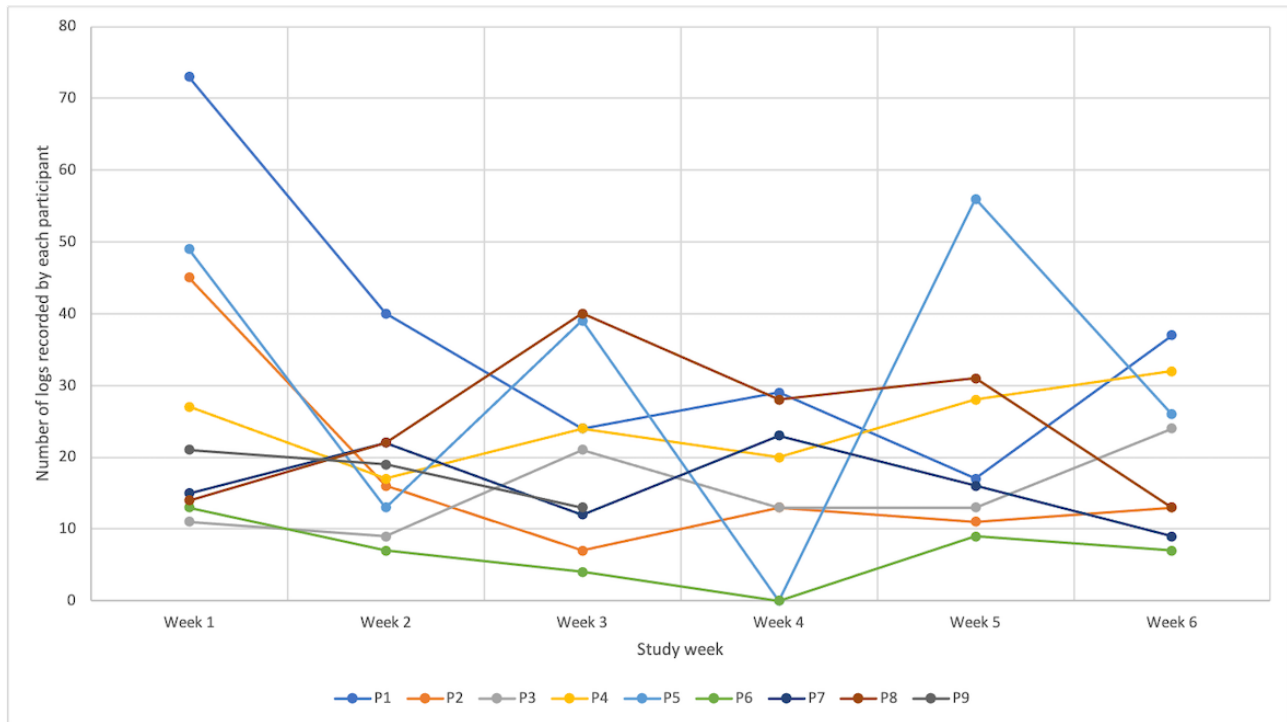


Table 4. Number of emotions logged through the prototypes.

Participants	Emotion Clock (n=579)	Emotion Board (n=506)
Participant #1 (n=220)	134	86
Participant #2 (n=105)	43	62
Participant #3 (n=91)	31	60
Participant #4 (n=148)	63	85
Participant #5 (n=183)	62	121
Participant #6 (n=40)	24	16
Participant #7 (n=97)	93	4
Participant #8 (n=148)	76	72
Participant #9 (n=53)	53	N/A ^a

^aNot applicable.

In examining the emotions that were logged, there are 3 main groupings, with neutral emotions being logged most frequently (relaxed [n=318], calm [n=276], and bored [n=72]), followed by positive emotions (happy [n=272] and excited [n=76]), with more negative emotions logged rarely (sad [n=31], nervous [n=26], and annoyed [n=14]).

Seven of our participants developed a routine as to when they logged emotions through the devices. Each of these routines was somewhat similar, with all of these participants regularly

logging in the morning and evenings, with additional logs throughout the day if seeing the device prompted them to think about logging:

...I have very low energy in the morning. So I usually changed it in the morning. I'd look at it usually, certainly, in the evening as well because at that point I'd be feeling more energetic and lively. During the day, I think, really if... Most of the time, I'm feeling fairly calm and cheerful. [Participant #1]

The remaining 2 participants had less of a routine around logging, relying on the prompts from the device or seeing the device as a reminder about logging:

...I don't think there was any specific time. It was when I suddenly thought, "Oh, I haven't done that yet," or I'd been out and I think, "I must do that when I get back." [Participant #2]

All but one of the participants placed the devices in the living room, perhaps the most public area in the home. This decision appeared to be driven by the convenience of using the device in the room most used and where the device would act as a reminder.

For the 8 participants who placed the devices in the living room, there was no concern about their last logged emotion being publicly visible, with their visitors generally being people they would freely discuss their emotions with (friends, family, etc). Instead, the devices acted as a talking point about the purpose of our study, which often led to a discussion of self-reflection:

...I found people were interested in it and often noticed it when they visited, and were interested in the whole idea. I had some friends round, there was quite a long conversation about mood and how you recognize mood. It was a talking point quite a lot of times... when you talked about it, they could recognize that it could be actually quite a clever way of getting you to recognize your mood and to understand how your mood changed. [Participant #1]

The participant who did not publicly display the devices, placed them within their study—a room they spend large amounts of time in (and were thus prompted by seeing the device), without advertising their emotions to visitors.

Perceived Need to Record Emotion

Five of our participants saw value in the devices as tools to monitor their own emotions, use that monitoring as a prompt for self-reflection and, if necessary, make changes to improve their emotional state:

...it's a good idea, because it makes you think about your mood, so therefore, you have to think before you select. So where, normally, I wouldn't bother-I'd just rush through the day. [Participant #5]

Eight of our participants also saw the monitoring as potentially a useful mechanism for sharing their feelings over time with other people. This was predominantly in the context of well-being and identifying whether family or friends needed to undertake some action as the person monitoring had seen a persistent or severe change in emotion. Of these 8 participants, 5 would have been happy to share their emotions with loved ones:

...I think I would be more open to indicating than saying probably. That might be just a man thing but it's you know I mean I just feel that I have to be happy and positive all the time. [Participant #7]

The remaining 3 participants indicated that they would be more comfortable with sharing with clinicians (eg, their doctor), would not be comfortable with sharing at all, or could see the

value in sharing but did not feel they were at that life stage yet (which did not correspond with participant age). This led us to consider whether participants who were less willing to share had a different profile of logged emotions (eg, whether they had a greater percentage of negative emotions). Comparing the participants' willingness to share their logged emotions with the emotions that participants had logged through the devices did not establish a clear pattern, with willingness to share more likely related to an individual's feelings of privacy.

Use of the Devices

Having noted that most participants identified a perceived need for the devices, it is necessary to consider what evidence there is that the devices had value to our participants. Five of our participants found that both of the devices helped them reflect on their emotions, with another 2 participants reporting this was only the case for the Emotion Clock and the Emotion Board. The ability to regularly log an emotion was a sufficient prompt to provide a scaffold for all of these participants to reflect on their emotional state:

...I think I thought about my moods quite a bit more, how I was really feeling, you know... Using it has had a positive impact, yes, because I've had to really think about how I feel. [Participant #2]

This was particularly the case during significant occasions. For participant #1 over their birthday and for participant #2 when their dog died, they found that the devices were particularly helpful in encouraging them to reflect on how they were feeling.

Most of our participants would like to continue using the devices. When explicitly asked whether they would like to continue monitoring their emotions using our devices, 5 of the participants saw clear value in them and would like to continue using them. None of these participants expressed a preference for only continuing with one of the devices. The remaining 4 participants did not like to continue using the devices, mainly as they did not perceive any derived benefit from their use. This included the 3 participants who did not consider themselves at a life stage of needing such a device; therefore, their disinterest was not a matter of dislike but rather of current lack in perceived need for emotional well-being management.

Device Preferences

Although the devices share certain design characteristics, the nature of interaction is significantly different. The clock offers a quick, immediate, and limited choice, while the board offers a more open-ended exploratory wide-ranging selection. It is worth examining how our participants engaged with these distinct designs and what can be learnt from those engagements.

For the Emotion Clock, 7 participants praised the simplicity of the design, stating:

"it was easy enough to use." [Participant #9]

These participants went on to discuss how the specificity of the emotions listed was not necessarily the emotions they wanted to record:

...I am, actually, a very busy person, which is why I say you should have that on there. If you're busy,

*you're not necessarily relaxed or calm. (Laughter)
You're just busy. Obviously, 'lonely' is not on here.*
[Participant #3]

This raises a question of the value of customizability, but in personalizing the words available for participants to select, the link between the device and the underlying psychological scales is removed. In contrast, only 2 participants felt that the Emotion Board (participant #3 and participant #4) was easy to use. Four participants felt that the Emotion Board was relatively difficult to understand, with the open nature of the interaction causing confusion:

...I sometimes found it a bit difficult to quite understand the square. I tended to move the thing round the edges of the square, I wasn't sure how the middle works and whether that calibrated things differently into the center. [Participant #1]

For some participants, this meant that they did not feel comfortable exploring the range of options through the Emotion Board, thereby reducing the use of the device as they did not understand the continuum nature of the design. However, 5 participants felt that while the Emotion Board was harder to understand, the necessary thought could help provoke further engagement and reflection:

...I had to think about that more... I certainly had to think about it more than with the [Emotion Clock], because it was whether you were feeling up, down, you know, energized, not energized. [Participant #3]

Participant #4 also noted that they related more to associating feelings with colors than they did with words, making the Emotion Board much more meaningful for them.

When our participants were asked which of the devices they preferred, the Emotion Clock was the most popular choice, with 6 of the participants preferring the simplicity of the interaction and the visual design. The other 2 participants, participant #4 and participant #8, preferred the open-ended interaction of the Emotion Board.

Design Characteristics

Having explored the specific design qualities of the individual devices, it is worth considering the design characteristics the devices shared and how they influenced our participants. The 2 devices shared certain design characteristics, particularly a shared aesthetics and a shared prompting system.

Six of our participants discussed the aesthetics of the devices without being prompted. All 6 were positive about the designs, noting that constructing the devices from wood made the devices pleasant to look at and made them blend in to the home environment. This is important as the aesthetics of the devices are likely an important factor as to whether people are likely to use the devices for long-term use; we would argue that if people are pleased by having the device in the house, they are much more likely to engage with the emotion logging in the long term.

As reported earlier, only 2 of our participants relied on the prompts for logging emotions, with the other 7 participants developing their own routine. All of the participants noted that the audio prompting was not annoying and not distracting.

Participant #1 noted that on occasion, the prompt could be useful as an occasional reminder, while participant #2 suggested increasing the frequency to 4 times a day as a more regular prompt. In general, though, our moderate prompting appears to have been appropriate.

Discussion

Value of the Devices

The focus of this work has been in evaluating the value of our tangible emotion logging devices for older adults. Our results demonstrate that our tangible devices can record data comparable to psychological scales of emotion. Such a finding validates the use of TUIs in this context and demonstrates that such devices could hold value for older adults. Furthermore, the level of use of the devices from our participants indicates that the participants saw some value in using the devices. The devices hold certain design properties that supported this use, particularly reflection on sparse data, provision of no data history, and focus on minimal interactions.

These properties are not unique in research into reflective logging technology. The value of reflecting on sparse data with minimal history is attracting increasing attention [70,71]. Further, focusing on minimal interaction is seen as a way for users to log meaningful data without becoming overburdened by the effort of logging [63,64,72]. We have built on this work and demonstrated that these design qualities in a different context—tangible devices for older adults—can support meaningful emotional reflection. Our findings open the design space for further consideration of how tangible devices can support emotional logging and reflection.

More specifically, our work also contributes to 2 ongoing interrelated debates within the field: the role of reflection in designs such as ours and the value of customizability in logging devices.

The Role of Reflection

Along with much of the human-computer interaction field, we have been somewhat imprecise in our treatment of reflection in our work, providing no firm definition or placing it within a theoretical framework [73]. To a certain extent, this was deliberate—our interest has been more on the design and success of the device rather than the mechanism through which users gained value. Although we operate under the assumption that the act of logging an emotional state would prompt users to think about their emotions and more broadly, their well-being in a form of reflection-in-action [58], we have not attempted to demonstrate that this mechanism is how our users gained value from the devices.

One of the key debates over supporting reflection through interaction design is the process by which reflection occurs. The model from Li et al [74] argues that reflection only happens at 1 stage of the reflection life-cycle, after preparation, collection, and integration, with the reflection leading to an action. This contrasts with the model from Epstein et al [61], which is more cyclical, with reflection taking place during an activity as well as afterwards.

Our research supports work that has demonstrated that people can reflect on relatively sparse data [75]. Our results suggest that a simple interaction, with no recorded history, is sufficient to support some users in reflecting on their emotional state. This is much closer to the Epstein et al's [61] model of reflection. None of our participants requested to see their recorded data at any point during the study, further suggesting that focusing on the design of the logging experience rather than on the historical record could be more beneficial to users.

One of the aims of personal informatics is to support behavior change and self-improvement by helping people become more self-aware. Some researchers have proposed that to do this effectively, we should not be constrained by supporting the consideration of past events but provide recommendations for future actions [76]. Such systems involve a combination of different subsystems. These include interfaces and device development, the design of analysis algorithms, and a complex sociotechnical mechanism for supporting the recommended actions.

Instead of attempting to construct all of the elements of such a system, we have focused on a single element (the interface design and device development), with results indicating that well-designed interfaces can be sufficient for some people to derive value from them. It remains an open question for the field as to whether such results can be enhanced by connecting such an interface to a well-designed and validated sociotechnical system for supporting deeper reflective actions. Given the complexity of the necessary "ongoing negotiation of the boundaries and meanings of self within an anxious alliance of knowledge, bodies, devices, and data" that is necessary for effective long-term use of logging technologies [77], we have provided a starting point for exploring the value of tangibles in this alliance.

The Value of Customizability

Some participants noted that they would have liked to have been able to customize the devices so that they were logging emotions more linked to their day-to-day experiences. Although this is perfectly feasible from a design perspective, it does remove the link between the device design and the underlying validated psychological scales being used. Our focus on ensuring the devices are linked to the validated psychological scales comes from the broader context of this work, where the research team is part of a project investigating home-based health monitoring technology. Working with clinicians, there was a focus on ensuring that if the data were later to be shared with clinicians or other stakeholders, it would be possible to understand the data in the context of an established framework.

This dichotomy is representative of a long-standing concern within the personal informatics community, with some researchers exploring better ways of aggregating and analyzing precise quantifiable data [78,79] and others arguing for a switch from a focus on "behavior and its objective data to the self and its subjective meanings" [71].

An alternative approach would be to design around affect labelling. This regulation technique can be described as asking people to put their feelings into words [80], which can help

people regulate their emotions [81]. This could prove an interesting route of customization for 2 reasons. First, it would be aggregating the labels in a meaningful way so that the historical record is useful to both the person logging and any related need (eg, with a clinician or carer). If the labels were restricted to a wide (but standardized) set such as Plutchik's Wheel of Emotions [82] or the Geneva Emotion Wheel [83], this aggregation could still take place automatically. Second, given the value of affect labelling comes from its open-ended nature, this is a design challenge in translating such a technique into a tangible logging tool.

Limitations and Further Work

We are working in an imprecise area of human experience. This means our findings and conclusions must be tempered by known limitations as discussed below.

Our first limitation stems from the design decisions we made. First, the Emotion Board makes strong use of color. Color is an inappropriate prompt for people with color blindness, and we have not accounted for the cultural implications inherent in color. Second, our devices do not cover fleeting emotions, as discussed by 2 of our participants. Third, by focusing on tangible technology suitable for the home, the resulting design was not suitable for logging emotions in outside contexts, as noted by 3 participants. Although we acknowledge these limitations as properties of our designs, they also indicate promising directions for further work.

The study methodology has a limitation in that we are unable to report the extent to which the participants' accuracy of interpreting the emotion expressed in the ANET vignettes was influenced by their personal ability to understand other people's emotions or their personal emotional reactions to the stimuli. We decided against screening participants based on their ability to interpret emotions from the vignettes and compensated for this by asking for the *participant descriptions*.

Additionally, we have no mechanism for comparing the data that participants logged during the field trial and how those participants were actually feeling. Although none of the participants raised this as an issue during the interviews, we cannot be completely certain as to whether participants tended to underlog or overlog particular types of emotions. Methodologically, this remains a challenge.

More broadly, our participant pool is relatively small and further work is needed to explore the generalizability of our results. The size of our study was directly limited by the COVID-19 pandemic, with one study cut short (participant #9) and 2 recruited participants unable to take part (participant #10, participant #11). Given that we were unable to safely distribute the tangible artefacts to a particularly COVID-vulnerable population, we were unable to extend the number of participants within the study. Furthermore, 6 of our participants were recruited through contact with the authors through word of mouth or previous participation in other studies. Although we have no personal relationship with these participants, they are more likely to be engaged in this kind of research and more technically able than the population as a whole. This convenience sampling also led to a gender imbalance among

our participants. While limiting the strength of the evidence, we are not arguing that our results are replicable across the population at large, but we argue that our work provides promising results and indicates further research directions.

Conclusion and Future Work

In this paper, we have contributed one of the first empirical investigations into the suitability of using tangible devices based on standardized scales of emotion for older adults to log emotions. We conclude that our devices are sufficiently accurate in collecting emotional data from older adults. Additionally, our work demonstrates the potential for using tangible devices to assist older adults in logging their emotional state to support reflection and emotional well-being. We argue that there is a significant amount of future work needed to extend this work by exploring whether this value holds when using tangibility as a design property of more self-expressive logging technology for older adults. Given the sharp divide between the competing

interests of generalizability and customizability, it is clear that designers have to establish what is more important to their user base. They should also ensure that their users have alternative options if their preferences change over time.

We argue that this success highlights the suitability for tangible devices to be used for long-term logging within the home. This study provides foundational support for tangible emotion self-logging devices for older adults and justifies further large-scale field studies exploring the effects of each device type on long-term engagement. In future work, we plan on exploring 2 interrelated aspects: (1) whether tangibility can be developed as a design quality for more self-expressive logging technologies and (2) exploring how to develop resilient sociotechnical support that responds to the data being logged by older adults. In doing so, we will better understand how tangible devices can help older adults wanting to maintain and improve their long-term well-being.

Acknowledgments

This work was completed under support from the UK Engineering and Physical Sciences Research Council grant EP/P01013X/1 and EP/V027263/1 as well as Science Foundation Ireland grant 13/RC/2094.

Conflicts of Interest

None declared.

Multimedia Appendix 1

Emotional logging summary tables.

[[PDF File \(Adobe PDF File\), 77 KB - humanfactors_v9i2e34606_app1.pdf](#)]

Multimedia Appendix 2

Interview questions.

[[PDF File \(Adobe PDF File\), 121 KB - humanfactors_v9i2e34606_app2.pdf](#)]

References

1. World population prospects: the 2017 revision. The United Nations. URL: https://www.un.org/development/desa/pd/sites/www.un.org.development.desa.pd/files/files/documents/2020/Jan/un_2017_world_population_prospects-2017_revision_databooklet.pdf [accessed 2022-04-11]
2. Mental health of older adults. World Health Organisation. 2017. URL: <https://www.who.int/news-room/fact-sheets/detail/mental-health-of-older-adults> [accessed 2022-04-11]
3. Future of an ageing population. UK Government Office for Science. URL: <https://www.gov.uk/government/publications/future-of-an-ageing-population> [accessed 2022-04-11]
4. Pynoos J, Nishita C, Kendig H. Housing. In: Birren JE, editor. Encyclopedia of Gerontology (Second Edition). New York: Elsevier; 2017:709-719.
5. Mynatt ED, Essa I, Rogers W. Increasing the opportunities for aging in place. 2000 Presented at: Proceedings on the 2000 conference on Universal Usability (CUU '00); November 16-17, 2000; Arlington, Virginia, USA. [doi: <https://doi.org/10.1145/355460.355475>]
6. Kendig H, Clemson L, Mackenzie L. Older people: well-being, housing and neighbourhoods. In: Smith SJ, editor. International Encyclopedia of Housing and Home. Amsterdam: Elsevier; 2012:150-155.
7. Wiles JL, Leibling A, Guberman N, Reeve J, Allen RES. The meaning of "aging in place" to older people. Gerontologist 2012 Jun;52(3):357-366. [doi: [10.1093/geront/gnr098](https://doi.org/10.1093/geront/gnr098)] [Medline: [21983126](https://pubmed.ncbi.nlm.nih.gov/21983126/)]
8. Kientz JA, Patel SN, Jones B, Price E, Mynatt ED, Abowd GD. The Georgia Tech aware home. 2008 Presented at: CHI '08 Extended Abstracts on Human Factors in Computing Systems; April 5-8, 2008; Montréal p. 3675-3680. [doi: [10.1145/1358628.1358911](https://doi.org/10.1145/1358628.1358911)]
9. Amiribesheli M, Benmansour A, Bouchachia A. A review of smart homes in healthcare. J Ambient Intell Human Comput 2015 Mar 14;6(4):495-517. [doi: [10.1007/s12652-015-0270-2](https://doi.org/10.1007/s12652-015-0270-2)]

10. Adib F, Mao H, Kabelac Z, Katabi D, Miller RC. Smart homes that monitor breathing and heart rate. 2015 Presented at: Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems; April 18-23, 2015; Seoul. [doi: [10.1145/2702123.2702200](https://doi.org/10.1145/2702123.2702200)]
11. A compendium of fact sheets: Well-being across the life course. UK Department of Health. 2014. URL: <https://www.gov.uk/government/publications/wellbeing-and-health-policy> [accessed 2022-04-11]
12. Massa P, Mazzali A, Zampini J, Zancanaro M. Quantify yourself: are older adults ready. In: Casiddu N, Porfirione C, Monteriù A, Cavallo F, editors. Ambient Assisted Living. Berlin/Heidelberg, Germany: Springer; 2019.
13. Doyle J, Walsh L, Sassu A, McDonagh T. Designing a wellness self-management tool for older adults: results from a field trial of YourWellness. 2014 Presented at: Proceedings of the 8th International Conference on Pervasive Computing Technologies for Healthcare; May 20-23, 2014; Brussels. [doi: [10.4108/icst.pervasivehealth.2014.254950](https://doi.org/10.4108/icst.pervasivehealth.2014.254950)]
14. Seiderer A, Hammer S, Andre E, Mayr M, Rist T. Exploring digital image frames for lifestyle intervention to improve well-being of older adults. 2015 Presented at: Proceedings of the 5th International Conference on Digital Health 2015; May 18, 2015; Florence, Italy. [doi: [10.1145/2750511.2750514](https://doi.org/10.1145/2750511.2750514)]
15. Desmet PM, Vastenburger MH, Romero N. Mood measurement with Pick-A-Mood: review of current methods and design of a pictorial self-report scale. *JDR* 2016;14(3):241. [doi: [10.1504/jdr.2016.079751](https://doi.org/10.1504/jdr.2016.079751)]
16. Broekens J, Brinkman W. AffectButton: A method for reliable and valid affective self-report. *International Journal of Human-Computer Studies* 2013 Jun;71(6):641-667 [FREE Full text] [doi: [10.1016/j.ijhcs.2013.02.003](https://doi.org/10.1016/j.ijhcs.2013.02.003)]
17. Sarzotti F. Self-Monitoring of Emotions and Mood Using a Tangible Approach. *Computers* 2018 Jan 08;7(1):7. [doi: [10.3390/computers7010007](https://doi.org/10.3390/computers7010007)]
18. Huisman G, van Hout M, van Dijk E, van der Geest T, Heylen D. LEMtool: measuring emotions in visual interfaces. 2013 Presented at: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems; April 27-May 3, 2013; Paris. [doi: [10.1145/2470654.2470706](https://doi.org/10.1145/2470654.2470706)]
19. Lee K, Hong H. Designing for self-tracking of emotion and experience with tangible modality. 2017 Presented at: Proceedings of the 2017 Conference on Designing Interactive Systems; June 10-14, 2017; Edinburgh. [doi: [10.1145/3064663.3064697](https://doi.org/10.1145/3064663.3064697)]
20. Rivera-Pelayo V, Fessl A, Müller L, Pammer V. Introducing Mood Self-Tracking at Work. *ACM Trans Comput-Hum Interact* 2017 Mar 22;24(1):1-28. [doi: [10.1145/3014058](https://doi.org/10.1145/3014058)]
21. Huang Y, Tang Y, Wang Y. Emotion map: a location-based mobile social system for improving emotion awareness and regulation. 2015 Presented at: Proceedings of the 18th ACM Conference on Computer Supported Cooperative Work & Social Computing; March 14-18, 2015; Vancouver. [doi: [10.1145/2675133.2675173](https://doi.org/10.1145/2675133.2675173)]
22. Hollis V, Konrad A, Springer A, Antoun M, Antoun C, Martin R, et al. What Does All This Data Mean for My Future Mood? Actionable Analytics and Targeted Reflection for Emotional Well-Being. *Human-Computer Interaction* 2017 Mar 16;32(5-6):208-267. [doi: [10.1080/07370024.2016.1277724](https://doi.org/10.1080/07370024.2016.1277724)]
23. Fuentes C, Herskovic V, Rodríguez I, Gereá C, Marques M, Rossel PO. A systematic literature review about technologies for self-reporting emotional information. *J Ambient Intell Human Comput* 2016 Nov 11;8(4):593-606. [doi: [10.1007/s12652-016-0430-z](https://doi.org/10.1007/s12652-016-0430-z)]
24. Doyle J, O'Mullane B, McGee S, Knapp RB. YourWellness: designing an application to support positive emotional well-being in older adults. 2012 Presented at: The 26th BCS Conference on Human Computer Interaction; September 12-14, 2012; Birmingham. [doi: [10.14236/ewic/hci2012.28](https://doi.org/10.14236/ewic/hci2012.28)]
25. Dasgupta D, Chaudhry B, Koh E, Chawla NV. A Survey of Tablet Applications for Promoting Successful Aging in Older Adults. *IEEE Access* 2016;4:9005-9017. [doi: [10.1109/access.2016.2632818](https://doi.org/10.1109/access.2016.2632818)]
26. Davidson JL, Jensen C. What health topics older adults want to track: a participatory design study. 2013 Presented at: Proceedings of the 15th International ACM SIGACCESS Conference on Computers and Accessibility; October 21-23, 2013; Bellevue, Washington. [doi: [10.1145/2513383.2513451](https://doi.org/10.1145/2513383.2513451)]
27. Ishii H. Tangible bits: beyond pixels. 2008 Presented at: Proceedings of the 2nd international conference on Tangible and embedded interaction; February 18-20, 2008; Bonn, Germany. [doi: [10.1145/1347390.1347392](https://doi.org/10.1145/1347390.1347392)]
28. Spreicer W. Tangible interfaces as a chance for higher technology acceptance by the elderly. 2011 Presented at: Proceedings of the 12th International Conference on Computer Systems and Technologies; June 16-17, 2011; Vienna, Austria. [doi: [10.1145/2023607.2023660](https://doi.org/10.1145/2023607.2023660)]
29. Suhas Govind J, Brathen H. Lowering the threshold: reconnecting elderly users with assistive technology through tangible interfaces. 2016 Presented at: International Conference on Human Aspects of IT for the Aged Population; July 17-22, 2016; Toronto. [doi: [10.1007/978-3-319-39943-0_6](https://doi.org/10.1007/978-3-319-39943-0_6)]
30. Bong WK, Chen W, Bergland A, Mandl T. Tangible User Interface for Social Interactions for the Elderly: A Review of Literature. *Advances in Human-Computer Interaction* 2018;2018:1-15 [FREE Full text] [doi: [10.1155/2018/7249378](https://doi.org/10.1155/2018/7249378)]
31. Gooch D, Mehta V, Price B, McCormick C, Bandara A, Bennaceur A, et al. How are you feeling? using tangibles to log the emotions of older adults. 2020 Presented at: Proceedings of the Fourteenth International Conference on Tangible, Embedded, and Embodied Interaction; February 9-12, 2020; Sydney, Australia. [doi: [10.1145/3374920.3374922](https://doi.org/10.1145/3374920.3374922)]
32. Frijda NH. Varieties of affect: emotions and episodes, moods, and sentiments. In: Ekman P, Davison R, editors. *The Nature of Emotions: Fundamental Questions*. New York: Oxford University Press; 1994.

33. Gross JJ. The Emerging Field of Emotion Regulation: An Integrative Review. *Review of General Psychology* 1998 Sep 01;2(3):271-299. [doi: [10.1037/1089-2680.2.3.271](https://doi.org/10.1037/1089-2680.2.3.271)]
34. Riva G, Villani D, Cipresso P, Repetto C, Triberti S, Di Lerna D, et al. Positive and Transformative Technologies for Active Ageing. *Stud Health Technol Inform* 2016;220:308-315. [Medline: [27046597](https://pubmed.ncbi.nlm.nih.gov/27046597/)]
35. Kanis M, Robben S, Kröse B. How Are You Doing? Enabling Older Adults to Enrich Sensor Data with Subjective Input. In *Proceedings of the 6th International Workshop on Human Behavior Understanding* 2015;9277:39-51. [doi: [10.1007/978-3-319-24195-1_4](https://doi.org/10.1007/978-3-319-24195-1_4)] [Medline: [11778978](https://pubmed.ncbi.nlm.nih.gov/11778978/)]
36. Pennebaker JW. Writing About Emotional Experiences as a Therapeutic Process. *Psychol Sci* 2016 May 06;8(3):162-166. [doi: [10.1111/j.1467-9280.1997.tb00403.x](https://doi.org/10.1111/j.1467-9280.1997.tb00403.x)]
37. Isaacs E, Konrad A, Walendowski A, Lennig T, Hollis V, Whittaker S. Echoes from the past: how technology mediated reflection improves well-being. 2013 Presented at: *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*; April 27-May 2, 2013; Paris. [doi: [10.1145/2470654.2466137](https://doi.org/10.1145/2470654.2466137)]
38. Hollis V, Konrad A, Whittaker S. Change of heart: emotion tracking to promote behavior change. 2015 Presented at: *Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems*; April 18-23, 2015; Seoul. [doi: [10.1145/2702123.2702196](https://doi.org/10.1145/2702123.2702196)]
39. Yang YS, Ryu GW, Choi M. Methodological Strategies for Ecological Momentary Assessment to Evaluate Mood and Stress in Adult Patients Using Mobile Phones: Systematic Review. *JMIR Mhealth Uhealth* 2019 Apr 01;7(4):e11215 [FREE Full text] [doi: [10.2196/11215](https://doi.org/10.2196/11215)] [Medline: [30932866](https://pubmed.ncbi.nlm.nih.gov/30932866/)]
40. Bradley MM, Lang PJ. Measuring emotion: The self-assessment manikin and the semantic differential. *Journal of Behavior Therapy and Experimental Psychiatry* 1994 Mar;25(1):49-59. [doi: [10.1016/0005-7916\(94\)90063-9](https://doi.org/10.1016/0005-7916(94)90063-9)]
41. Russell JA. A circumplex model of affect. *Journal of Personality and Social Psychology* 1980;39(6):1161-1178. [doi: [10.1037/h0077714](https://doi.org/10.1037/h0077714)]
42. Barrett LF, Russell JA. The Structure of Current Affect. *Curr Dir Psychol Sci* 2016 Jun 22;8(1):10-14. [doi: [10.1111/1467-8721.00003](https://doi.org/10.1111/1467-8721.00003)]
43. Mehrabian A, Russell JA. *An Approach to Environmental Psychology*. Cambridge: MIT Press; 1974.
44. Read JC. Validating the Fun Toolkit: an instrument for measuring children's opinions of technology. *Cogn Tech Work* 2007 May 22;10(2):119-128. [doi: [10.1007/s10111-007-0069-9](https://doi.org/10.1007/s10111-007-0069-9)]
45. Goodday SM, Atkinson L, Goodwin G, Saunders K, South M, Mackay C, et al. The True Colours Remote Symptom Monitoring System: A Decade of Evolution. *J Med Internet Res* 2020 Jan 15;22(1):e15188 [FREE Full text] [doi: [10.2196/15188](https://doi.org/10.2196/15188)] [Medline: [31939746](https://pubmed.ncbi.nlm.nih.gov/31939746/)]
46. Tsanas A, Saunders KEA, Bilderbeck AC, Palmius N, Osipov M, Clifford GD, et al. Daily longitudinal self-monitoring of mood variability in bipolar disorder and borderline personality disorder. *J Affect Disord* 2016 Nov 15;205:225-233 [FREE Full text] [doi: [10.1016/j.jad.2016.06.065](https://doi.org/10.1016/j.jad.2016.06.065)] [Medline: [27449555](https://pubmed.ncbi.nlm.nih.gov/27449555/)]
47. Perez Arribas I, Goodwin GM, Geddes JR, Lyons T, Saunders KEA. A signature-based machine learning model for distinguishing bipolar disorder and borderline personality disorder. *Transl Psychiatry* 2018 Dec 13;8(1):274 [FREE Full text] [doi: [10.1038/s41398-018-0334-0](https://doi.org/10.1038/s41398-018-0334-0)] [Medline: [30546013](https://pubmed.ncbi.nlm.nih.gov/30546013/)]
48. Palmius N, Osipov M, Bilderbeck AC, Goodwin GM, Saunders K, Tsanas A, et al. A multi-sensor monitoring system for objective mental health management in resource constrained environments. 2014 Presented at: *Appropriate Healthcare Technologies for Low Resource Settings*; September 17-18, 2014; London. [doi: [10.1049/cp.2014.0764](https://doi.org/10.1049/cp.2014.0764)]
49. Zhang R, Nicholas J, Knapp AA, Graham AK, Gray E, Kwasny MJ, et al. Clinically Meaningful Use of Mental Health Apps and its Effects on Depression: Mixed Methods Study. *J Med Internet Res* 2019 Dec 20;21(12):e15644 [FREE Full text] [doi: [10.2196/15644](https://doi.org/10.2196/15644)] [Medline: [31859682](https://pubmed.ncbi.nlm.nih.gov/31859682/)]
50. Chandrashekar P. Do mental health mobile apps work: evidence and recommendations for designing high-efficacy mental health mobile apps. *Mhealth* 2018;4:6 [FREE Full text] [doi: [10.21037/mhealth.2018.03.02](https://doi.org/10.21037/mhealth.2018.03.02)] [Medline: [29682510](https://pubmed.ncbi.nlm.nih.gov/29682510/)]
51. Morris ME, Kathawala Q, Leen TK, Gorenstein EE, Guilak F, Labhard M, et al. Mobile therapy: case study evaluations of a cell phone application for emotional self-awareness. *J Med Internet Res* 2010 Apr 30;12(2):e10 [FREE Full text] [doi: [10.2196/jmir.1371](https://doi.org/10.2196/jmir.1371)] [Medline: [20439251](https://pubmed.ncbi.nlm.nih.gov/20439251/)]
52. Fernández M, Rodríguez I, Rossel PO, Fuentes C. InMyDay: a digital diary to promote self-care among elders. 2017 Presented at: *11th International Conference on Ubiquitous Computing & Ambient Intelligence*; November 7-10, 2017; Philadelphia. [doi: [10.1007/978-3-319-67585-5_49](https://doi.org/10.1007/978-3-319-67585-5_49)]
53. Fuentes C, Rodríguez I, Herskovic V. EmoBall: A Study on a Tangible Interface to Self-report Emotional Information Considering Digital Competences. 2015 Presented at: *Ambient Intelligence for Health*; December 1-4, 2015; Puerto Varas p. 189-200. [doi: [10.1007/978-3-319-26508-7_19](https://doi.org/10.1007/978-3-319-26508-7_19)]
54. Balaam M, Fitzpatrick G, Good J, Luckin R. Exploring affective technologies for the classroom with the subtle stone. 2010 Presented at: *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*; April 10-15, 2010; Atlanta. [doi: [10.1145/1753326.1753568](https://doi.org/10.1145/1753326.1753568)]
55. Jingar M, Lindgren H. Tangible communication of emotions with a digital companion for managing stress: an exploratory co-design study. 2019 Presented at: *Proceedings of the 7th International Conference on Human-Agent Interaction*; October 6-10, 2019; Kyoto. [doi: [10.1145/3349537.3351907](https://doi.org/10.1145/3349537.3351907)]

56. State of musculoskeletal health report. Arthritis Research UK. URL: <https://www.versusarthritis.org/about-arthritis/data-and-statistics/the-state-of-musculoskeletal-health/>
57. Salivia G, Hourcade JP. PointAssist: assisting individuals with motor impairments. 2013 Presented at: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems; April 27-May 2, 2013; Paris. [doi: [10.1145/2470654.2466157](https://doi.org/10.1145/2470654.2466157)]
58. Schon D. The Reflective Practitioner: How Professionals Think In Action. New York: Basic Books; 1984.
59. Ayobi A, Sonne T, Marshall P, Cox AL. Flexible and mindful self-tracking: design implications from paper bullet journals. 2018 Presented at: Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems; April 21-26, 2018; Montreal. [doi: [10.1145/3173574.3173602](https://doi.org/10.1145/3173574.3173602)]
60. Choe EK, Abdullah S, Rabbi M, Thomaz E, Epstein DA, Cordeiro F, et al. Semi-Automated Tracking: A Balanced Approach for Self-Monitoring Applications. IEEE Pervasive Comput 2017 Jan;16(1):74-84. [doi: [10.1109/mprv.2017.18](https://doi.org/10.1109/mprv.2017.18)]
61. Epstein DA, Ping AN, Fogarty J, Munson SA. A lived informatics model of personal informatics. 2015 Presented at: Proceedings of the 2015 ACM International Joint Conference on Pervasive and Ubiquitous Computing; September 7-11, 2015; Osaka. [doi: [10.1145/2750858.2804250](https://doi.org/10.1145/2750858.2804250)]
62. Thudt A, Hinrichs U, Huron S, Carpendale S. Self-reflection and personal physicalization construction. 2018 Presented at: Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems; April 21-26, 2018; Montreal. [doi: [10.1145/3173574.3173728](https://doi.org/10.1145/3173574.3173728)]
63. Choe EK, Lee B, Kay M, Pratt W, Kientz JA. SleepTight: low-burden, self-monitoring technology for capturing and reflecting on sleep behaviors. 2015 Presented at: Proceedings of the 2015 ACM International Joint Conference on Pervasive and Ubiquitous Computing; September 7-11, 2015; Osaka. [doi: [10.1145/2750858.2804266](https://doi.org/10.1145/2750858.2804266)]
64. Ferrario MA, Simm W, Gradinar A, Forshaw S, Smith MT, Lee T, et al. Computing and mental health: intentionality and reflection at the click of a button. 2017 Presented at: Proceedings of the 11th EAI International Conference on Pervasive Computing Technologies for Healthcare; May 23-26, 2017; Barcelona. [doi: [10.1145/3154862.3154877](https://doi.org/10.1145/3154862.3154877)]
65. Lazar A, Koehler C, Tanenbaum TJ, Nguyen DH. Why we use and abandon smart devices. 2015 Presented at: Proceedings of the 2015 ACM International Joint Conference on Pervasive and Ubiquitous Computing; September 7-11, 2015; Osaka. [doi: [10.1145/2750858.2804288](https://doi.org/10.1145/2750858.2804288)]
66. Bradley MM, Lang PJ. Affective norms for English text (ANET): affective ratings of text and instruction manual. The Center for the Study of Emotion and Attention. 2017. URL: <https://csea.phhp.ufl.edu/media/anetmessage.html> [accessed 2022-04-11]
67. Lombard M, Snyder-Duch J, Bracken CC. Content Analysis in Mass Communication: Assessment and Reporting of Intercoder Reliability. Human Comm Res 2002 Oct;28(4):587-604. [doi: [10.1111/j.1468-2958.2002.tb00826.x](https://doi.org/10.1111/j.1468-2958.2002.tb00826.x)]
68. Braun V, Clarke V. Using thematic analysis in psychology. Qualitative Research in Psychology 2006 Jan;3(2):77-101. [doi: [10.1191/1478088706qp063oa](https://doi.org/10.1191/1478088706qp063oa)]
69. Landis JR, Koch GG. The Measurement of Observer Agreement for Categorical Data. Biometrics 1977 Mar;33(1):159. [doi: [10.2307/2529310](https://doi.org/10.2307/2529310)]
70. Elsdon C, Selby M, Durrant A, Kirk D. Fitter, happier, more productive. Interactions 2016 Aug 23;23(5):45-45. [doi: [10.1145/2975388](https://doi.org/10.1145/2975388)]
71. Rapp A, Tirassa M. Know Thyself: A Theory of the Self for Personal Informatics. Human-Computer Interaction 2017 Apr 10;32(5-6):335-380. [doi: [10.1080/07370024.2017.1285704](https://doi.org/10.1080/07370024.2017.1285704)]
72. Chong MK, Whittle J, Rashid U, Ang CS. Cue now, reflect later: a study of delayed reflection of diary events. 2015 Presented at: Human-Computer Interaction-INTERACT; September 14-18, 2015; Bamberg. [doi: [10.1007/978-3-319-22698-9_24](https://doi.org/10.1007/978-3-319-22698-9_24)]
73. Baumer EPS, Khovanskaya V, Matthews M, Reynolds L, Sosik VS, Gay G. Reviewing reflection: on the use of reflection in interactive system design. 2014 Presented at: Proceedings of the 2014 conference on Designing interactive systems; June 21-25, 2014; Vancouver. [doi: [10.1145/2598510.2598598](https://doi.org/10.1145/2598510.2598598)]
74. Li I, Dey A, Forlizzi J. A stage-based model of personal informatics systems. 2010 Presented at: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems; April 10-15, 2010; Atlanta. [doi: [10.1145/1753326.1753409](https://doi.org/10.1145/1753326.1753409)]
75. Whooley M, Ploderer B, Gray K. On the integration of self-tracking data amongst quantified self members. 2014 Presented at: Proceedings of the 28th International BCS Human Computer Interaction Conference; September 9-12, 2014; Southport. [doi: [10.14236/ewic/hci2014.19](https://doi.org/10.14236/ewic/hci2014.19)]
76. Hollis V, Konrad A, Springer A, Antoun M, Antoun C, Martin R, et al. What Does All This Data Mean for My Future Mood? Actionable Analytics and Targeted Reflection for Emotional Well-Being. Human-Computer Interaction 2017 Mar 16;32(5-6):208-267. [doi: [10.1080/07370024.2016.1277724](https://doi.org/10.1080/07370024.2016.1277724)]
77. Williams K. An Anxious Alliance. AAHCC 2015 Oct 05;1(1):11. [doi: [10.7146/aaahcc.v1i1.21146](https://doi.org/10.7146/aaahcc.v1i1.21146)]
78. Bentley F, Tollmar K, Stephenson P, Levy L, Jones B, Robertson S, et al. Health Mashups. ACM Trans. Comput.-Hum. Interact 2013 Nov 01;20(5):1-27. [doi: [10.1145/2503823](https://doi.org/10.1145/2503823)]
79. Jones SL, Kelly R. Dealing With Information Overload in Multifaceted Personal Informatics Systems. Human-Computer Interaction 2017 May 16;33(1):1-48. [doi: [10.1080/07370024.2017.1302334](https://doi.org/10.1080/07370024.2017.1302334)]
80. Lieberman MD, Eisenberger NI, Crockett MJ, Tom SM, Pfeifer JH, Way BM. Putting feelings into words: affect labeling disrupts amygdala activity in response to affective stimuli. Psychol Sci 2007 May;18(5):421-428. [doi: [10.1111/j.1467-9280.2007.01916.x](https://doi.org/10.1111/j.1467-9280.2007.01916.x)] [Medline: [17576282](https://pubmed.ncbi.nlm.nih.gov/17576282/)]

81. Torre JB, Lieberman MD. Putting Feelings Into Words: Affect Labeling as Implicit Emotion Regulation. *Emotion Review* 2018 Mar 20;10(2):116-124. [doi: [10.1177/1754073917742706](https://doi.org/10.1177/1754073917742706)]
82. Plutchik R. The Nature of Emotions. *Am Sci* 2001;89(4):344. [doi: [10.1511/2001.4.344](https://doi.org/10.1511/2001.4.344)]
83. Scherer KR. What are emotions? And how can they be measured? *Social Science Information* 2016 Jun 29;44(4):695-729. [doi: [10.1177/0539018405058216](https://doi.org/10.1177/0539018405058216)]

Abbreviations

ANET: Affective Norms for English Text

SAM: self-assessment manikin

TUI: tangible user interface

Edited by A Kushniruk; submitted 01.11.21; peer-reviewed by A Joseph, E Stringer; comments to author 28.11.21; revised version received 07.01.22; accepted 06.03.22; published 27.04.22.

Please cite as:

Gooch D, Mehta V, Stuart A, Katz D, Bennasar M, Levine M, Bandara A, Nuseibeh B, Bennaceur A, Price B

Designing Tangibles to Support Emotion Logging for Older Adults: Development and Usability Study

JMIR Hum Factors 2022;9(2):e34606

URL: <https://humanfactors.jmir.org/2022/2/e34606>

doi: [10.2196/34606](https://doi.org/10.2196/34606)

PMID: [35475781](https://pubmed.ncbi.nlm.nih.gov/35475781/)

©Daniel Gooch, Vikram Mehta, Avelie Stuart, Dmitri Katz, Mohamed Bennasar, Mark Levine, Arosha Bandara, Bashar Nuseibeh, Amel Bennaceur, Blaine Price. Originally published in *JMIR Human Factors* (<https://humanfactors.jmir.org>), 27.04.2022. This is an open-access article distributed under the terms of the Creative Commons Attribution License (<https://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in *JMIR Human Factors*, is properly cited. The complete bibliographic information, a link to the original publication on <https://humanfactors.jmir.org>, as well as this copyright and license information must be included.

Original Paper

Determinants of Laypersons' Trust in Medical Decision Aids: Randomized Controlled Trial

Marvin Kopka^{1,2*}, MSc; Malte L Schmieding^{1*}, MD; Tobias Rieger³, MSc; Eileen Roesler³, MSc; Felix Balzer^{1*}, MSc, MD, PhD; Markus A Feufel^{4*}, Dipl-Ing (FH), MSc, PhD

¹Institute of Medical Informatics, Charité – Universitätsmedizin Berlin, corporate member of Freie Universität Berlin and Humboldt-Universität zu Berlin, Berlin, Germany

²Cognitive Psychology and Ergonomics, Department of Psychology and Ergonomics (IPA), Technische Universität Berlin, Berlin, Germany

³Work, Engineering and Organizational Psychology, Department of Psychology and Ergonomics (IPA), Technische Universität Berlin, Berlin, Germany

⁴Division of Ergonomics, Department of Psychology and Ergonomics (IPA), Technische Universität Berlin, Berlin, Germany

*these authors contributed equally

Corresponding Author:

Marvin Kopka, MSc

Institute of Medical Informatics

Charité – Universitätsmedizin Berlin, corporate member of Freie Universität Berlin and Humboldt-Universität zu Berlin

Charitépl 1

Berlin, 10117

Germany

Phone: 49 30 450 581 052

Email: marvin.kopka@charite.de

Abstract

Background: Symptom checker apps are patient-facing decision support systems aimed at providing advice to laypersons on whether, where, and how to seek health care (disposition advice). Such advice can improve laypersons' self-assessment and ultimately improve medical outcomes. Past research has mainly focused on the accuracy of symptom checker apps' suggestions. To support decision-making, such apps need to provide not only accurate but also trustworthy advice. To date, only few studies have addressed the question of the extent to which laypersons trust symptom checker app advice or the factors that moderate their trust. Studies on general decision support systems have shown that framing automated systems (anthropomorphic or emphasizing expertise), for example, by using icons symbolizing artificial intelligence (AI), affects users' trust.

Objective: This study aims to identify the factors influencing laypersons' trust in the advice provided by symptom checker apps. Primarily, we investigated whether designs using anthropomorphic framing or framing the app as an AI increases users' trust compared with no such framing.

Methods: Through a web-based survey, we recruited 494 US residents with no professional medical training. The participants had to first appraise the urgency of a fictitious patient description (case vignette). Subsequently, a decision aid (mock symptom checker app) provided disposition advice contradicting the participants' appraisal, and they had to subsequently reappraise the vignette. Participants were randomized into 3 groups: 2 experimental groups using visual framing (anthropomorphic, 160/494, 32.4%, vs AI, 161/494, 32.6%) and a neutral group without such framing (173/494, 35%).

Results: Most participants (384/494, 77.7%) followed the decision aid's advice, regardless of its urgency level. Neither anthropomorphic framing (odds ratio 1.120, 95% CI 0.664-1.897) nor framing as AI (odds ratio 0.942, 95% CI 0.565-1.570) increased behavioral or subjective trust ($P=.99$) compared with the no-frame condition. Even participants who were extremely certain in their own decisions (ie, 100% certain) commonly changed it in favor of the symptom checker's advice (19/34, 56%). Propensity to trust and eHealth literacy were associated with increased subjective trust in the symptom checker (propensity to trust $b=0.25$; eHealth literacy $b=0.2$), whereas sociodemographic variables showed no such link with either subjective or behavioral trust.

Conclusions: Contrary to our expectation, neither the anthropomorphic framing nor the emphasis on AI increased trust in symptom checker advice compared with that of a neutral control condition. However, independent of the interface, most participants trusted the mock app's advice, even when they were very certain of their own assessment. Thus, the question arises as to whether laypersons use such symptom checkers as substitutes rather than as aids in their own decision-making. With trust in symptom

checkers already high at baseline, the benefit of symptom checkers depends on interface designs that enable users to adequately calibrate their trust levels during usage.

Trial Registration: Deutsches Register Klinischer Studien DRKS00028561; <https://tinyurl.com/rv4utcfb> (retrospectively registered).

(*JMIR Hum Factors* 2022;9(2):e35219) doi:[10.2196/35219](https://doi.org/10.2196/35219)

KEYWORDS

symptom checkers; disposition advice; anthropomorphism; artificial intelligence; urgency assessment; patient-centered care; human-computer interaction; consumer health; information technology; IT; mobile phone

Introduction

Background

Patients are increasingly searching for health information on the web before seeking medical care [1-3]. As an alternative to commercial search engines, patient-facing decision support systems called symptom checkers were developed to provide the first access point to health-related information. These tools are targeted at laypersons and ask users to enter their signs and symptoms before presenting preliminary diagnoses and an assessment of the level of care to seek [4]. The latter assessment, the so-called disposition or urgency advice, is arguably the more important function of symptom checkers, as it could prevent unnecessary visits and direct patients toward the appropriate health care facility, thus reducing the burden on the health care system [5,6].

Related Work

Symptom checkers have mostly been investigated in terms of accuracy; user characteristics; and, occasionally, their effect on user care-seeking behavior. We will report on these findings in turn. In non-industry-funded studies, their *accuracy* appears to be mediocre: Semigran et al [6] found that disposition advice of apps is accurate at 57% on average, Yu et al [7] identified an accuracy between 50% and 74% for emergency cases, and Hill et al [8] found appropriate disposition advice to be provided in 49% of case evaluations on average. Although the symptom checker accuracy in these studies is mediocre, the range is very broad, and some symptom checkers perform well. For example, Ceney et al [9] found a disposition accuracy of up to 90% for a system that performs best in urgency assessment. At the health system level, evidence is still inconclusive whether symptom checkers bear the potential to make patient journeys more efficient and decrease the burden on health care services, with a study on telephone triage suggesting a redistribution rather than a reduction in health care workload [4,10-12]. Given that the current reliability of symptom checkers seems rather low on average, 2 (systematic) reviews advise against using these tools in lieu of current assessment models [13,14].

Concerning user characteristics, research has found that symptom checker *users* are predominantly female, more often young than old, and more often have a higher than a lower level of education [15,16]. In terms of *behavioral effects*, one study showed that most users plan to follow the received advice [17]. Another study by Winn et al [18] found that the perceived urgency of symptoms decreased after using a symptom checker. However, the advice given by the symptom checker was not

recorded in that study, and it remains unclear whether users are more prone to lower urgency advice or whether they might have overestimated the urgency in their initial assessment. A vignette-based experimental study found that on average, symptom checkers currently do not outperform laypersons in terms of disposition accuracy. However, best-in-class apps seem superior to laypersons [19,20].

In addition to a system's accuracy, which is well known to affect behavior, subjective trust (ie, self-reported trust in automated systems; see the study by Schaefer et al [21]) is a key factor determining whether humans follow advice from decision aids or rely on automated systems (behavioral trust). Trust in automation has been shown to be influenced by several factors, which can be divided into performance based (eg, reliability) and attribute based (eg, appearance) [21]. Although symptom checker research has so far focused on performance-based factors, studies on the influence of attribute-based factors are mostly missing. In general automation research, anthropomorphism—making the automation appear human like—has been identified as one of many potential influences [22]. There are several methods for designing human-like systems or framing them as such; however, visual anthropomorphism is the easiest to include in a symptom checker (eg, using a picture of a person on the user interface). The direction of the relationship between visual anthropomorphism and trust seems to vary. In a study by de Visser et al [23], trust was lower for anthropomorphic interfaces compared with technical systems. However, this is only the case when the system's reliability is high. With decreasing reliability, trust decreased less steeply for the anthropomorphic system than for the technical system, suggesting a resilient influence of anthropomorphism, which could be replicated in another study [24].

In contrast, in a medical decision-making task, Pak et al [25] found that trust and follow rates, with constant reliability of 67%, were higher when the decision support system's interface included the image of a physician. These contradictory findings might be explained by Hertz and Wiese [26], who found that people preferred assistive agents that were thought to have the greatest expertise for a specific task. For medical decision-making, health care professionals are highly trusted, and patients seem to ascribe greater expertise to physicians than to self-assessment apps, whereas in other use cases, such as analytical and computational tasks, users might find assistance from a nonhuman agent more trustworthy [1,27]. In terms of symptom checkers, anthropomorphic framing could be used to increase expertise perception because of humanization (ie,

making it more human like) or technological framing (eg, artificial intelligence [AI]) because of technologization (ie, emphasizing its technological nature so that it is seen as an expert system). Indeed, some symptom checkers, such as Symptomate [28], have already emphasized using AI algorithms, which are commonly used as a buzzword for machines imitating human intelligence [29]. Although transparent communication of using AI in applications will soon be required by law [30], the design of such systems often hints at displaying AI use to enhance trust because of increased expertise perception [31]. On the basis of these findings, will showing an image of a physician in symptom checkers make them more trustworthy? Could trust also be enhanced by emphasizing that symptom checkers base their recommendations on AI?

Aim of This Study

This study aimed to examine the influence of framing effects on subjective trust in symptom checkers and the behavioral consequences of trust (ie, dependence and following behavior), which are strongly related [32]. Higher trust is particularly useful when using highly accurate symptom checkers, as patient outcomes can only be improved by following correct (and safe) advice. However, when a symptom checker does not perform well, high trust can also be dangerous in the case of incorrect advice (eg, recommending self-care while emergency care is required). Thus, our study aimed to identify potential factors influencing users' trust in these decision aids apart from system accuracy.

As trust in physicians is generally higher than in computerized decision aids [27], we were particularly interested in assessing whether anthropomorphic framing (ie, displaying an image of a physician as a human expert decision maker on the user interface) leads to increased trust in decision aids. Furthermore, we examined whether framing the symptom checker as being based on AI increases users' trust. We hypothesized that anthropomorphism would increase participants' subjective trust in the app and the proportion of participants following the app's advice (behavioral trust). We expected the same effect (higher subjective and behavioral trust) when framing the symptom checker as AI. As Winn et al [18] showed that users commonly decreased their appraised urgency level after symptom checker use, we explored whether users might be more prone to follow a symptom checker when its urgency appraisal is lower than their own. Kopka and colleagues [33] found that most laypersons are certain in their urgency assessment and that in absolute numbers, laypersons make most errors when they are certain of their appraisal. For this reason, we also examined whether users tended to accept advice from such decision aids when they were already certain of their own judgment. We expected that users' inclination to follow a decision aid's advice would decrease with higher decisional certainty, as users tend to rely on automation when they are not confident but solve tasks manually when they are confident [34,35]. Finally, we explored the association between demographical and other interindividual variables and trust.

Methods

Ethics Approval and Consent to Participate

This study was approved by the Ethics Committee of the Department of Psychology and Ergonomics (Institut für Psychologie und Arbeitswissenschaft [IPA]) at Technische Universität Berlin (tracking number: FEU_9_210315). Participants volunteered to participate in the survey, and informed consent was required. On the first page, participants were told about the investigator, the study's purpose, what data were to be collected during the study, and where and for how long they would be stored. On the second page, participants were informed about the duration of the survey (approximately 5 minutes) and received additional information regarding the scope and use of attention checks.

Participants

Yee et al [36] found that the effect size for showing a human in an interface on subjective trust was Cohen's $d=0.28$. On the basis of an a priori power analysis for independent t tests with an assumed Cronbach α of .05 and a power of $1-\beta=0.80$, we aimed to sample at least 477 ($n=159$, $n=159$, and $n=159$ for the 3 groups, respectively) participants to detect differences between the 3 groups (2 experimental groups and 1 control group). We expected some participants to fail attention checks (items that were embedded in the survey questions and asked participants to select a particular option, eg, "Please select Disagree"); therefore, we oversampled by 10%. To avoid participants' decisions being influenced by their residential country or ability to understand the scenario, only US residents fluent in English were eligible. They also had to participate in the web-based questionnaire on a desktop device or tablet as the survey's graphical elements could not be reliably displayed on smartphone devices. Another requirement was not being a medical professional (ie, nurse, paramedic, and physician). We sampled participants using Prolific [37], a platform characterized by high data quality [38], starting on Saturday, May 15, 2021, at 5 PM Eastern Daylight Time and on Sunday, May 16, 2021, at 4 PM Eastern Daylight Time. We chose these days as Casey et al [39] have shown that the samples recruited via the web are more diverse during the weekend than on working days. Following Ho et al [40], participants were compensated £0.70 (US \$0.91) for their participation and received an additional £0.18 (US \$0.24) as an incentive for the correct decision (ie, selecting self-care in their last appraisal) to increase data quality through attentive participation.

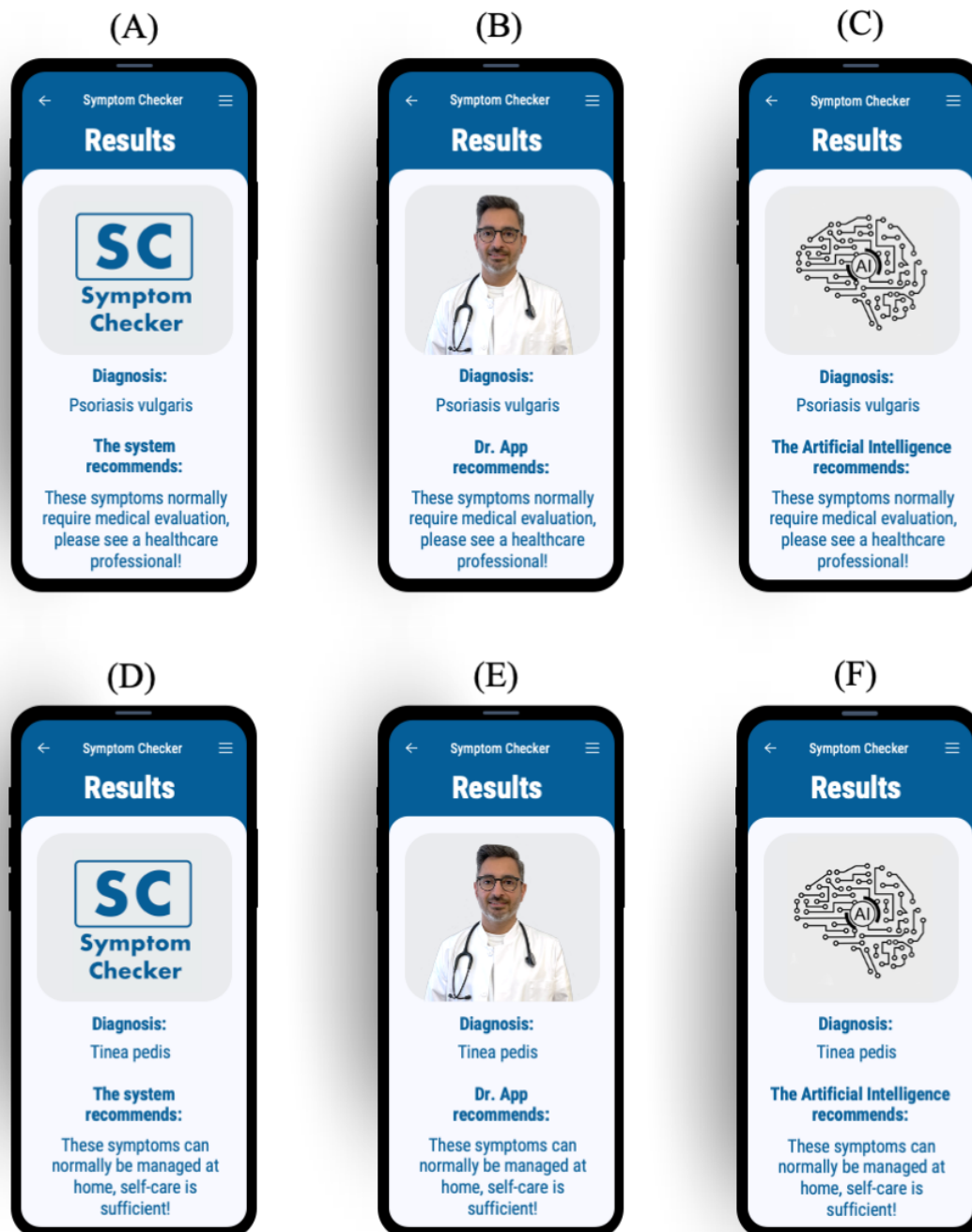
Design

We used a 1-factorial experimental design with factor framing and factor levels of anthropomorphic framing and framing as AI along with a control group (allocation ratio 1:1:1). These were manipulated by integrating a picture of a physician, an iconographic representation of AI similar to that displayed by Symptomate [28], or a mock company logo into the mock symptom checker's advice screen (Figure 1). Participants were automatically randomly assigned (simple randomization) to one of these levels using a randomization tool integrated into the Unipark Enterprise Feedback Suite (EFS) Survey. In every condition, they had to appraise one and the same case vignette

by deciding whether the fictitious patient required health care or self-care was sufficient. Although there are other urgency levels in symptom checkers, we chose this binary decision as the question of whether to seek care at all is the first decision patients must make [15,20]. The participants were tasked to appraise the case vignette twice: once before receiving advice

from the decision aid (initial stand-alone assessment) and once after receiving the advice. The decision aid's advice was programmed to always contradict the participant's stand-alone assessment. The dependent variables were subjective and behavioral trust (ie, whether the participant followed the advice of the decision aid).

Figure 1. Interfaces: participants were asked about their initial appraisal and received contrary advice from the results screen of a mock symptom checker. Screens (A), (B), and (C) present advice to seek health care, whereas screens (D), (E), and (F) recommend that self-care is sufficient. Participants were randomized and received advice from a neutral (A) and (D), anthropomorphic (B) and (E), or artificial intelligence-framed (C) and (F) screen.



Materials

We chose one specific case vignette, originally used by Hill et al [8] to assess symptom checker accuracy, as it showed a high

item variance; that is, participants in an unpublished pretest with 56 participants were about equally split in their appraisal of whether the fictitious patient required health care. Owing to the high item variance, this case vignette allowed us to capture

the potential influence of the advice's urgency (ie, higher or lower urgency) on users' trust. The case vignette describes the typical presentation of fungal skin infection (tinea pedis), colloquially known as athlete's foot: "A 33-year-old male has scaly skin between the toes. The skin is a little itchy and turns soft and white when wetted. The skin has an odour." According to the gold standard solution for this case vignette assigned by an expert panel [8], proper self-care was sufficient for this case. Although not necessary, we would also consider it appropriate to seek professional health care for the condition, as a physician could educate the patient on proper self-care options. Hence, it is neither negligent in regarding self-care as sufficient nor is it overcautious to deem the health care required when appraising this vignette. As the decision aid always disagreed with the participants' initial assessment, it either gave higher or lower urgency advice depending on the participant's initial stand-alone assessment. As most symptom checkers complement their disposition recommendation with a diagnostic assessment, our decision aid named a possible diagnosis congruent with the corresponding urgency level: when providing disposition advice more acute than the participant (ie, the symptom checker recommended seeking health care), the symptom checker provided the (made-up) diagnosis of psoriasis vulgaris along with the text, "These symptoms normally require medical evaluation, please see a healthcare professional!" For lower urgency, it returned the (original) diagnostic suggestion of tinea pedis along with the text, "These symptoms can normally be managed at home, self-care is sufficient!" The wording is based on a screening of different symptom checkers and represents a symbiosis between the advice given by Symptomate [28] and Ada [41].

As a decision aid, we created a mock app with a simple result presentation screen using PowerPoint (Microsoft Corporation) [42], Affinity Photo (Serif Ltd) [43], and Vectorator (Linearity) [44]. Participants could not interact with the decision aid to input information as not everyone would have entered the symptoms in the same way, and thus, the decision path would have differed. As this interaction influences trust [45], we tried to eliminate any resulting bias by presenting a results screen only. This design was inspired by Pak et al [25], who assessed anthropomorphism in a decision support system for diabetics using the picture of a physician. We designed our interfaces to include the same diagnoses and disposition advice with a picture of a mock symptom checker logo, a physician, or an icon representing AI (Figure 1). To ensure that the decision aid was displayed in the same way for all participants and that the results were not biased by different presentations on different phones, we placed the interface directly in a mock phone. The simulated phone could then be viewed on a computer or tablet. For the anthropomorphic condition, we chose the depiction of a young male physician based on the findings of a study by Pak et al [46], who found the depictions of a male physician embedded in a decision aid less susceptible to fluctuations in perceptions of trust as a function of the decision aid's reliability, and depictions of younger agents exhibited fewer age differences in perceived trust than older agents.

Framing manipulation corresponds to the actual visual framing in a widely used symptom checker [28]. Although framing can

be manipulated to a greater extent (eg, by presenting videos and stories [23]), we decided to use a picture only to represent currently applied practice. Therefore, the extent of our manipulation is similar to that of other studies that assessed the effects of anthropomorphism on trust in decision aids [25,47].

Survey

A web-based survey was developed using Unipark EFS Survey [48]. All collected data were saved on the platform, and only the authors of this study had access to the data. We evaluated the usability and technical functionality of the questionnaire and then conducted a pilot study in which test participants were asked to provide feedback on any display problems, unclear questions or statements, or other issues that might have occurred. After these were resolved, the questionnaire was rolled out as a voluntary, open survey that was only accessible via the Prolific recruitment platform (initial contact). We did not advertise the survey in any other way than presenting it on the platform. Participants were presented with 1 questionnaire on each page; hence, the items per page ranged from 1 to 13 on a sum of 19 pages. They could return using browser buttons and review their answers, which were checked for completeness using the built-in function of the Unipark EFS Survey. Although the symptom checker interfaces were adapted to the participant's responses (see the *Design* section and the *Materials* section), we did not use adaptive questioning to reduce the number of questions.

Survey visitor numbers were assessed by assigning participants an ID when opening the questionnaire. Most participants accessing the survey completed it (completion rate: 572/607, 94.2%).

Dependent Measures

Subjective trust in the symptom checker app (primary outcome) was measured by adapting the Trust in Automated Systems Survey [49], which uses a 7-point Likert scale with 12 items; as suggested by Gutzwiller et al [50], we randomized the order in which items were presented to avoid a positive bias. Behavioral trust (secondary outcome) was measured using an adapted TNO trust task [51], as previously reported by several authors [23,24,47,52]. First, the participants had to rate the appropriate urgency level on their own. Afterward, they were shown the symptom checker app's contradicting recommendation and had to make a final decision. We measured whether they changed their decision in favor of the decision aid's advice and coded behavioral trust at the individual level as Boolean (true or false). We then determined the proportion of participants following the advice (follow rates) as a measure of behavioral trust at the group level.

Procedure

After participants gave consent to participate, we surveyed their age, gender, educational background, and prior medical training. Next, participants were asked about their propensity to trust using the Propensity to Trust in Technology Scale with 6 items on a 5-point Likert scale [53] and their eHealth literacy using the eHealth Literacy Scale with 8 items on a 5-point Likert scale [54].

Afterward, the 2 urgency levels (*health care* and *self-care*) between which the participants had to choose when appraising the fictitious case vignette were explained, and participants' understanding of these definitions was assured with multiple-choice questions (3 rewordings of the urgency level definitions to which participants had to assign the correct urgency level). Next, they were shown the case vignette, and they appraised its urgency and rated their decisional certainty using a visual analogue scale with values from 0 (minimum certainty) to 100 (maximum certainty). They then saw the results screen of the mock decision aid advising the opposite of their assessment with 1 of the 3 different designs (Figure 1) and had to make a final decision on the urgency level and again state their decisional certainty.

Thereafter, they were presented with the Trust in Automated Systems Survey and had the opportunity to provide feedback or any other comments in an open text field. Finally, to ensure that our intervention was successful, participants were asked to specify which image was embedded in the decision aid presented to them previously (manipulation check).

Data Analysis

We cleaned and analyzed the data using base R (version 4.0.5) [55], the *tidyverse* packages [56], and *aod* [57]. For inferential analysis of continuous outcomes, we used a 1-way between-subjects ANOVA. For binary outcomes, we used dummy-coded binomial logistic regression and tested the coefficients using Wald chi-square tests. To test demographic and interindividual influences, we used multiple linear regression and multiple binomial logistic regression with

standardized coefficients for better comparability. The effect coding scheme and results can be found in [Multimedia Appendices 1-6](#). We used an effect coding scheme to compare each factor level to the mean of all factor levels. Thus, deviations from the mean can be quantified and tested for significance instead of performing group comparisons with a single, consistent reference category (as in dummy coding). For example, the coefficients and *P* values of gender 1 ([Multimedia Appendix 1](#)) represent the differences and significance tests of women compared with the mean of other genders. Similarly, the metrics of education 1 ([Multimedia Appendix 2](#)) represent differences between participants with a bachelor's degree compared with the mean of all other education levels. Finally, we conducted sensitivity power analyses using the R package *pwr* [58] to estimate the population effect size for selected results that appeared statistically nonsignificant.

Results

Participant Characteristics

The survey was completed in 6 minutes and 19 seconds (Median, IQR 4 minutes, 36 seconds to 8 minutes, 46 seconds). Of the 607 individuals accessing the survey, 35 (5.8%) did not finish the questionnaire, 14 (2.3%) were excluded as they were trained medical professionals, 27 (4.4%) were excluded as they took part on a mobile phone, and 37 (6.1%) failed at least one of the attention checks. Therefore, of the 607 individuals, the total sample size was 494 (81.4%). Distributions of age, gender, level of education, propensity to trust, and eHealth literacy overall and in each of the 3 groups are reported in [Table 1](#).

Table 1. Participant characteristics (N=494).

Characteristics	Control group (n=173)	Anthropomorphic (n=160)	Artificial intelligence (n=161)	Total
Age (years), mean (SD)	34.5 (13.8)	32.1 (12.5)	31.6 (12.2)	32.8 (12.9)
Gender, n (%)				
Female	81 (46.8)	78 (48.8)	77 (47.8)	236 (47.8)
Male	87 (50.3)	80 (50)	82 (50.9)	249 (50.4)
Other	5 (2.9)	2 (1.3)	2 (1.2)	9 (1.8)
Education, n (%)				
Less than high school	0 (0)	4 (2.5)	3 (1.9)	7 (1.4)
High school graduate	25 (14.5)	12 (7.5)	20 (12.4)	57 (11.5)
College or associate degree	48 (27.7)	50 (31.3)	63 (39.1)	161 (32.6)
Bachelor degree	66 (38.2)	66 (41.3)	52 (32.3)	184 (37.2)
Graduate degree or higher	34 (19.7)	28 (17.5)	23 (14.3)	85 (17.2)
Prior medical training, n (%)				
No training	141 (81.5)	135 (84.4)	136 (84.3)	412 (83.4)
Basic first aid	32 (18.5)	25 (15.6)	25 (15.5)	82 (16.6)
Propensity to Trust score ^a , mean (SD)	4.1 (0.5)	4.1 (0.5)	4.0 (0.5)	4.1 (0.5)
eHEALS ^b score ^c , mean (SD)	30.5 (4.91)	30.0 (5.27)	30.1 (5.61)	30.2 (5.25)
Initial assessment of the case vignette, n (%)				
Health care	60 (34.7)	62 (38.8)	65 (40.4)	187 (37.9)
Self-care	113 (65.3)	98 (61.3)	96 (59.6)	307 (62.1)
Completion time (minutes), median (IQR)	6:32 (4:32-8:53)	6:09 (4:39-8:17)	6:18 (4:38-9:16)	6:19 (4:36-8:46)

^aPropensity to Trust refers to the Propensity to Trust in Technology Scale, and possible scores range from 1 (low) to 5 (high).

^beHEALS: eHealth Literacy Scale.

^cPossible scores range from 8 (low) to 40 (high).

Almost all participants (480/494, 97.2%) correctly recollected the image embedded in the mock decision aid (manipulation check).

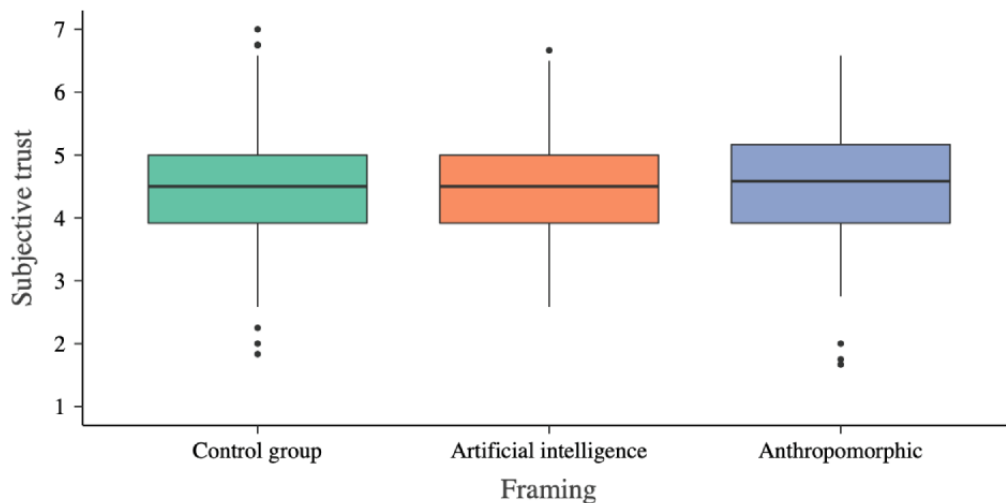
Influence of Framing on Participants' Trust

Subjective Trust

Descriptively, trust in all 3 framing conditions was very similar (anthropomorphic: mean 4.503, SD 0.922; AI: mean 4.495, SD

0.817; control: mean 4.508, SD 0.921; [Figure 2](#)). Framing had no significant effect on subjective trust ($F_{2,491}=0.009$; $P=.99$; $\eta^2=0.00$). On the basis of a sensitivity power analysis ($\alpha=.05$; $1-\beta=0.80$; anthropomorphic $n=160$, AI $n=173$, and control $n=161$), we estimated the effect size of possible differences between the groups to not be greater than $\eta^2=0.018$.

Figure 2. Subjective trust scores across the 3 study groups. Trust was operationalized using the Trust in Automated Systems Survey with a range from 1 (minimum trust) to 7 (maximum trust). The horizontal line in the box represents the median.



Behavioral Trust

Most participants followed the decision aid’s advice and changed their urgency appraisal (384/494, 77.7%). Behavioral trust was slightly higher for the anthropomorphic system (127/160, 79.4%) than for the control group (134/173, 77.5%); however, the difference (odds ratio [OR] 1.120, 95% CI 0.664-1.897) was not statistically significant ($\chi^2_1=0.2$; $P=.67$).

Behavioral trust was slightly lower for the AI system (123/161, 76.4%) than for the control group; however, the difference (OR 0.942, 95% CI 0.565-1.570) was not statistically significant ($\chi^2_1=0.053$; $P=.82$), either.

Influence of the Urgency Level Provided by the Decision Aid on Trust

We observed no differences in subjective trust between participants receiving advice of greater urgency (*health care required*) than their stand-alone initial assessment (*self-care sufficient*) (mean 4.5, SD 0.919) and those receiving less urgent advice (mean 4.5, SD 0.869). Concerning behavioral trust, the

proportion of participants who followed more urgent advice was slightly lower (235/307, 76.5%) than the proportion who followed advice of lower urgency than their own initial stand-alone assessment (149/187, 79.7%).

Influence of Participants’ Decisional Certainty on Trust

The participants of all 3 groups were certain about their initial stand-alone assessment (median 70, IQR 60-81). Only 12.8% (63/494) were unsure (ie, indicating a certainty of <50% about their appraisal). No differences in patterns were observed between the framing conditions.

Participants’ certainty in their initial assessment was not associated with subjective trust in the decision aids ($R^2=0.001$; Figure 3). With increasing decisional certainty, behavioral trust decreased (OR 0.966, 95% CI 0.952-0.979; $\chi^2_1=25.0$; $P<.001$; McFadden $R^2=0.055$). However, behavioral trust was high and remained >50% (19/34, 56%), even for participants indicating maximum decision certainty (100/100; Figure 4). There were no differences between the framing conditions.

Figure 3. Subjective trust and participants’ certainty. Trust was operationalized using the Trust in Automated Systems Survey (range: 1-7). The dashed blue indicates a linear model for the association between participants’ certainty in their initial stand-alone appraisal of the case vignette and the subjective trust toward the decision aid.

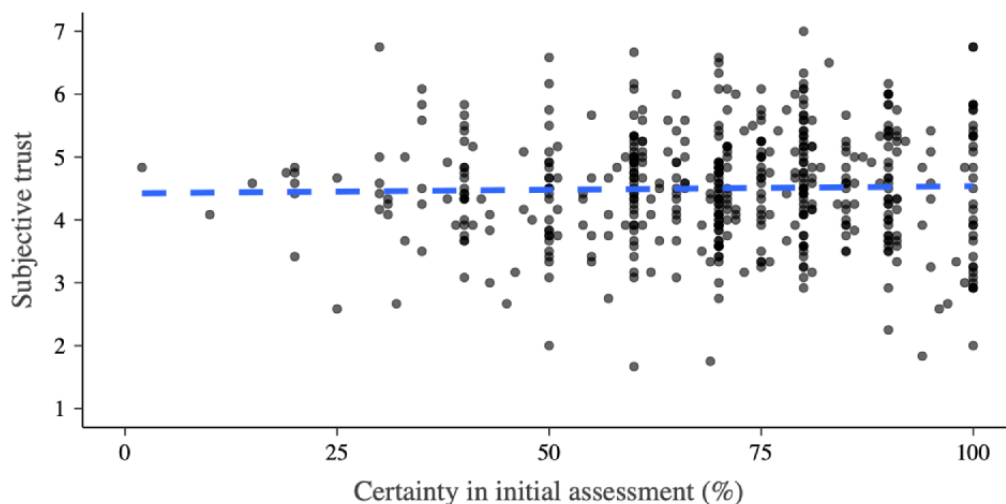
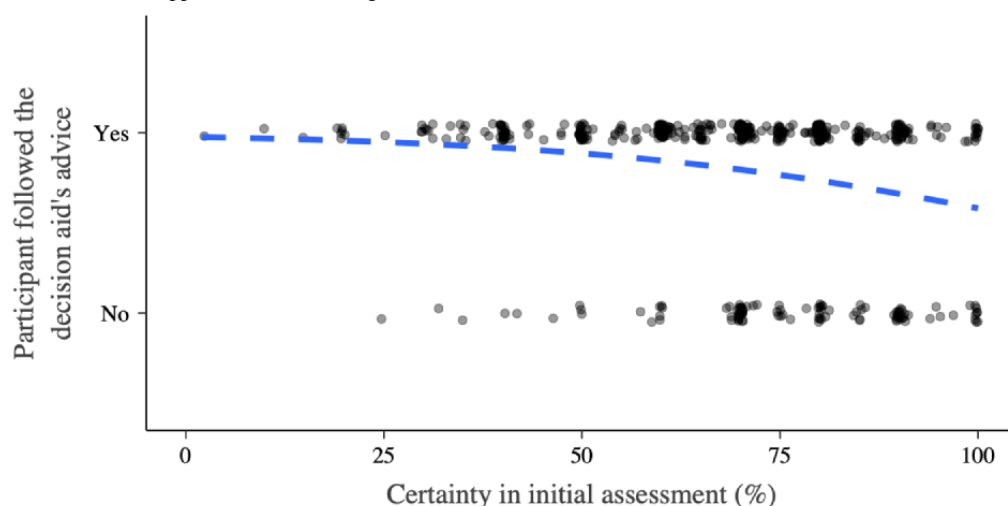


Figure 4. Behavioral trust and participants' certainty. The dashed blue indicates a binomial logistic model for the association between participants' certainty in their initial stand-alone appraisal of the case vignette and the behavioral trust toward the decision aid.



Demographic and Interindividual Influences on Trust

Neither demographic variables (age, gender, and education) nor basic first aid training was associated with subjective and behavioral trust in the symptom checker when controlling for the other variables. However, an individual's propensity for trust and eHealth literacy increased subjective trust and were statistically significant ($P < .001$). However, these 2 variables did not have a statistically significant influence on behavioral trust (Multimedia Appendices 4 and 6).

Discussion

Principal Findings

Effect of Anthropomorphic or AI Framing

The aim of this study was to explore the factors influencing laypersons' subjective and behavioral trust in symptom checkers. In particular, we examined the hypothesis that the common features of symptom checker interfaces that frame the system as either AI-based or anthropomorphic affect users' trust in these systems. Our analysis does not support this hypothesis: we could not observe a difference in trust—neither subjective nor behavioral—between a neutral symptom checker interface (showing a mock company logo) and interfaces framed as either anthropomorphic or as using AI. This is in contrast to previous findings from other domains where anthropomorphism led to an increase or decrease in trust [23–25]. In addition, we expected that designing anthropomorphic decision aids in a clinical context would yield higher trust as symptom checker users trust physicians more than self-assessment apps [15,26,27,59]. However, our participants' trust was unaffected by how the symptom checker was framed, indicating that users seem to perceive a symptom checker mainly as an app, regardless of whether a depiction of a physician or an AI icon is included. We did not find an effect of framing on subjective and behavioral trust in our study; however, we cannot rule out that framing might influence other variables in the trust formation process. For example, anthropomorphism has been shown to moderate the relationship between reliability and trust [23,24]: framing might moderate the impact of a decision aid's reliability

on trust; however, it might not be sufficient to build trust (in medical advice) on its own. Instead, other factors, such as explanations of the reasoning underlying the symptom checker's advice, might help build trust more effectively [60].

Persuasive Power of Symptom Checker Apps

We found that most participants (384/494, 77.7%) followed the decision aid's advice. This is in line with Verzantvoort et al [17], who reported a high intention of users to follow dispositional advice from a decision aid (65%). However, both findings stand in contrast to those indicating a low behavioral trust in symptom checkers; for example, Meyer et al [15] reported that only a minority of those advised by a symptom checker to visit the emergency department followed this advice, and Miller et al [61] found that most patients presenting to a primary care clinic stick to their stand-alone assessment when using a symptom checker in a primary care clinic's waiting room. Taken together, these findings hint at symptom checker users' behavioral trust being a function of the exact urgency decision and context of use: when users are undecided between seeking emergency or nonemergency care, they might depend less on the symptom checker's advice compared with when choosing whether professional medical care is required at all or self-care is appropriate. Users might also be more inclined to accept guiding advice from symptom checker apps before arriving at a health care facility. Interestingly, a web search seems to change only few people's urgency level [62]; this indicates a difference in advice-taking between symptom checker use and general web search.

Another influencing factor on behavioral trust is decisional certainty; that is, whether users follow a symptom checker's advice depends on how certain they are of their own stand-alone assessment. However, our findings hint at the high persuasive power of symptom checkers: although participants indicating maximum certainty in their own stand-alone assessment followed the advice less often than those indicating lower levels of certainty, most still changed their decision according to the decision aid's recommendation. This finding is central as it emphasizes the impact symptom checkers may have on the urgency decision. Symptom checkers could not only assist when

patients are uncertain whether and where to seek health care but also convince those who (wrongly) are very certain in their appraisal. This might prove very useful, as Kopka and colleagues [33] report that laypersons' urgency errors are most frequent when they indicate high confidence in their stand-alone appraisal. In contrast, high dependence on symptom checkers potentially signifies laypersons using them as a replacement for decision-making rather than as a decision aid. This should be further investigated through research on the cognitive and metacognitive mechanisms with which laypersons monitor their own reasoning when confronted with advice from symptom checkers, similar to the Jussupow et al [63] study on an AI-based decision aid supporting physicians in diagnostic decisions. In addition, laypersons' high dependence on symptom checkers emphasizes the need for a framework to identify and label those apps defying the general trend by proving them to be both accurate and safe to use, as currently, symptom checkers' accuracy is being reported as mediocre in general, with only a few performing well [6-9,14].

Interindividual Variables' Effect on Trust in Symptom Checkers

A previous study indicated gender differences in appraising medical situations (eg, Cooper and Humphrey [64] showed that female participants assessed their urgency as more risk averse); however, we could not replicate this finding for trust. Our findings suggest that demographic and interindividual differences might be negligible when drafting recommendations on whether and how symptom checker apps should be designed. Although users who are generally more inclined to trust show higher subjective trust in symptom checkers, they do not seem to follow their advice more often, which might have methodological reasons; that is, the item terminology referencing technology too broadly [53,65]. Concerning the influence of eHealth literacy on trust, we observed that it increased subjective trust but not behavioral trust. Users with higher eHealth literacy might have more knowledge about eHealth applications and thus be more open to receiving advice from a decision aid while at the same time being more able to integrate a decision aid's advice into their own decision-making rather than uncritically adopting the presented advice.

Limitations

First, the intervention might not have been effective in producing meaningful differences. However, nearly all participants (480/494, 97.2%) were able to recall the picture they were presented with as part of the decision aid, thus proving that they took note of the depictions used for framing. Moreover, the results remained consistent, even if participants who could not recall the presented picture were excluded from the analysis. The framing itself represents another limitation. Although we followed the current practice and manipulation extent of previous studies, other interface and framing aspects are conceivable that may not have been captured in this study. For example, it would be interesting to assess whether personalized images (eg, patients' own physicians) could increase their trust.

In our study, participants did not interact with the decision aid as we only presented a symptom checker's results screen instead of letting them enter the data or symptoms into an actual app.

This was done to keep the survey short, avoid dropout when entering symptoms for a longer period, and avoid introducing any bias because of different algorithmic pathways resulting from participants unreliably entering information, which is a nonnegligible risk, as shown by Jungmann et al [66]. As trust could be influenced by user experience throughout the interaction [45], we could not account for a potentially moderating role of that factor. This limitation applies equally to all experimental groups; thus, internal validity is not compromised. However, as symptom checkers commonly require extensive user interaction over a span of multiple minutes [67], their ecological validity might be limited. Future research should alter the existing symptom checkers to test whether our results can be replicated in practice. Our participants also only evaluated a single case vignette, whereas, in other studies, participants solved as many as 20 with the help of a symptom checker app. Hence, the duration of exposure to the intervention was low in our study. However, we consider this closer to the real use case of symptom checkers, where users seek advice on a single set of complaints rather than systematically testing the app by iteratively entering the signs and symptoms of highly heterogeneous fictitious patient descriptions. However, unlike in the real use case, participants could not change their decision at a later stage. In practice, they might decide to see a health care professional after gathering further evidence, even if they decided for self-care to be sufficient when using a symptom checker. Thus, our concept of behavioral trust only captures users' intentions after consulting a symptom checker, not their actual behavior (ie, [not] seeking health care according to the symptom checker's prompt).

All participants appraised only a single case vignette, which was the same across all 3 groups. We used only this specific case vignette as it has been used in previous studies and was ambiguous enough for patients to choose both self-care and health care. However, technically, many other vignettes and symptoms can be entered and should thus be investigated in the future. The gold standard for the case vignette used in this study is self-care; however, visiting a health care professional with these symptoms is not inappropriate and, in particular, not unsafe. Thus, deviation from the gold standard solution may be considered wrong but not consequential. Although the gold standard solution was assigned by a panel of physicians, the idea of absolute correct urgency may vary for different physicians. It would be interesting to see whether our findings can be replicated for a variety of cases with different gold standard urgency levels (eg, 3-tiered or 4-tiered urgency levels). Other decisions, such as whether emergency care is required, should also be examined, as this study could not provide any evidence for other urgency-related decisions. Especially concerning the decision of whether emergency care is required, we consider a further investigation into the question of whether layperson trust is unaffected by the direction of the (contradicting) advice by a decision aid worthwhile, as here, an incorrect appraisal is more consequential.

It cannot be ruled out that some participants researched the correct solution on the web to obtain a bonus. However, as this

could have occurred in all groups, internal validity should not be impaired.

Participants did not assess their own symptoms but were presented with a fictitious case vignette as a proxy for a medical case. Although this arguably reductionist approach is commonly applied when evaluating symptom checkers [6,8,68,69], it remains unclear whether participants assess these symptoms in the same way they do when experiencing them. For example, in the case of real symptoms, not only might the information input change, but the patients' mental well-being and their perceived self-efficacy in implementing an action might also have an impact. It is also conceivable that participants might not have empathized enough with the situation or that the urgency was assessed differently. However, web-based health information sources are commonly used to assess the symptoms of others [3]; thus, this use case still possesses a high degree of external validity.

As we only collected quantitative data, we could not explain why the participants changed their decisions. Future studies should conduct qualitative studies on decision-making when assisted by a symptom checker.

Finally, the participants in this study were well-educated, with 54.5% (269/494) of participants having a bachelor's degree or higher. Although our sample is not representative of the US population, the average education level is very close to that of symptom checker users [15]. The same applies to our participants' average age, which is very close to that of users [16] and had no impact on our exploratory analyses.

Practical Implications

Although some developers frame their symptom checkers as anthropomorphic or as an AI, there appears to be no meaningful impact on users' trust based on our study. Although previous studies found an influence of anthropomorphism on trust in general automation [23-25], we could not extend these findings to symptom checkers. As we kept this study as true to reality as possible—by specifically using a mock symptom checker instead of other decision aids used in experimental laboratory setups and by testing an externally valid use case where users only assessed a single case vignette and could not estimate symptom checker accuracy—our results are more applicable to the specific use case of symptom checkers. The effect found by

other authors presumably materializes only when users can assess a system's accuracy. As multiple assessments in a row do not correspond to the natural use of symptom checkers, framing (as currently applied) does not seem to provide any benefit in terms of trust.

Although sociodemographic factors appear to have an impact on symptom checker use [15,16], they do not seem to alter trust. Thus, trust depends on eHealth literacy; for example, symptom checkers do not need to be customized for age, gender, or education to increase trust, although they might be customized to increase usability and user experience.

Finally, as initial trust is very high, regardless of framing and demographic factors, further increasing users' trust in these systems may not be a priority. Instead, we suggest that it may be more worthwhile to explore ways of supporting users in their decision-making so that they do not have to rely uncritically on a symptom checker's advice. For example, this can be achieved by providing explanations of disposition advice tailored to the individual user [60].

Conclusions

The subjective and behavioral trust of laypersons in clinical decision aids is high and is not influenced by framing such systems as anthropomorphic or using AI.

However, users are inclined to change their minds based on the symptom checker's advice, even when they report maximum certainty in their initial and contradicting stand-alone appraisal. This indicates the high persuasive power of the symptom checker and thus demonstrates its potential to make patient journeys more efficient. At the same time, our findings hint at the danger that laypersons may use symptom checkers to substitute rather than to assist their own decision-making. Although some symptom checkers commonly provide accurate and safe advice, the range of symptom checker accuracy varies widely. Thus, before recommending symptom checkers for general use, rigorous standards for evaluating symptom checkers must be defined to ensure that only those symptom checkers are recommended that are accurate and safe enough to be worthy of the trust people have in them. Further research should investigate how to ensure that symptom checkers function as aids rather than replacements in laypersons' decision-making.

Acknowledgments

The authors express their gratitude to the participants and Sami Gaber for the portrait pictures. This project was funded by the home institutions of the last authors (MAF and FB). The authors acknowledge the support from the German Research Foundation (Deutsche Forschungsgemeinschaft [DFG]) and the Open Access Publication Fund of Charité, Universitätsmedizin Berlin.

Authors' Contributions

MK and MLS conceived the study. MK created the questionnaire and interfaces, designed and conducted the analyses, and wrote the first draft of the manuscript. MLS assisted in questionnaire creation and data analysis and worked on manuscript development. ER, TR, FB, and MAF provided critical input and advice on the study and questionnaire design, analysis methods, and drafts of the paper. All authors accept full responsibility for the final version of this manuscript.

Conflicts of Interest

All authors have completed the International Committee of Medical Journal Editors uniform disclosure form and declare no support from any organization for the submitted work, no financial relationships with any organizations that might have an interest in the submitted work in the previous 3 years, and no other relationships or activities that could appear to have influenced the submitted work. The lead author affirms that this manuscript is an honest, accurate, and transparent account of the study being reported; that no important aspects of the study have been omitted; and that any discrepancies from the study as planned (and, if relevant, registered) have been explained.

Editorial Notice

This randomized study was only retrospectively registered, given that the authors believed registration was unnecessary for this kind of trial and the particular outcomes which are measured. The editor granted an exception from ICMJE rules mandating prospective registration of randomized trials. However, readers are advised to carefully assess the validity of any potential explicit or implicit claims related to primary outcomes or effectiveness, as retrospective registration does not prevent authors from changing their outcome measures retrospectively.

Multimedia Appendix 1

Effect coding scheme of gender.

[[DOCX File , 13 KB - humanfactors_v9i2e35219_app1.docx](#)]

Multimedia Appendix 2

Effect coding scheme of education.

[[DOCX File , 13 KB - humanfactors_v9i2e35219_app2.docx](#)]

Multimedia Appendix 3

Multiple linear regression of demographic and interindividual influences on subjective trust with standardized coefficients.

[[DOCX File , 14 KB - humanfactors_v9i2e35219_app3.docx](#)]

Multimedia Appendix 4

Multiple binomial logistic regression of demographic and interindividual influences on behavioral trust with standardized coefficients.

[[DOCX File , 14 KB - humanfactors_v9i2e35219_app4.docx](#)]

Multimedia Appendix 5

Multiple linear regression of demographic and interindividual influences on subjective trust with unstandardized coefficients.

[[DOCX File , 14 KB - humanfactors_v9i2e35219_app5.docx](#)]

Multimedia Appendix 6

Multiple binomial logistic regression of demographic and interindividual influences on behavioral trust with unstandardized coefficients.

[[DOCX File , 14 KB - humanfactors_v9i2e35219_app6.docx](#)]

Multimedia Appendix 7

CONSORT-eHEALTH checklist (V 1.6.1).

[[PDF File \(Adobe PDF File\), 323 KB - humanfactors_v9i2e35219_app7.pdf](#)]

References

1. Baumann E, Czerwinski F, Rosset M, Seelig M, Suhr R. Wie informieren sich die Menschen in Deutschland zum Thema Gesundheit? Erkenntnisse aus der ersten Welle von HINTS Germany. Bundesgesundheitsblatt Gesundheitsforschung Gesundheitsschutz 2020 Sep;63(9):1151-1160. [doi: [10.1007/s00103-020-03192-x](https://doi.org/10.1007/s00103-020-03192-x)] [Medline: [32666180](https://pubmed.ncbi.nlm.nih.gov/32666180/)]
2. Cocco AM, Zordan R, Taylor DM, Weiland TJ, Dilley SJ, Kant J, et al. Dr Google in the ED: searching for online health information by adult emergency department patients. Med J Aust 2018 Oct 15;209(8):342-347. [doi: [10.5694/mja17.00889](https://doi.org/10.5694/mja17.00889)] [Medline: [30107763](https://pubmed.ncbi.nlm.nih.gov/30107763/)]
3. Fox S, Duggan M. Health Online 2013. Pew Internet & American Life Project. 2013 Jan 15. URL: https://www.pewinternet.org/wp-content/uploads/sites/9/media/Files/Reports/PIP_HealthOnline.pdf [accessed 2021-07-13]
4. Chambers D, Cantrell AJ, Johnson M, Preston L, Baxter SK, Booth A, et al. Digital and online symptom checkers and health assessment/triage services for urgent health problems: systematic review. BMJ Open 2019 Aug 01;9(8):e027743 [FREE Full text] [doi: [10.1136/bmjopen-2018-027743](https://doi.org/10.1136/bmjopen-2018-027743)] [Medline: [31375610](https://pubmed.ncbi.nlm.nih.gov/31375610/)]

5. Millenson ML, Baldwin JL, Zipperer L, Singh H. Beyond Dr. Google: the evidence on consumer-facing digital tools for diagnosis. *Diagnosis (Berl)* 2018 Sep 25;5(3):95-105 [[FREE Full text](#)] [doi: [10.1515/dx-2018-0009](https://doi.org/10.1515/dx-2018-0009)] [Medline: [30032130](https://pubmed.ncbi.nlm.nih.gov/30032130/)]
6. Semigran HL, Linder JA, Gidengil C, Mehrotra A. Evaluation of symptom checkers for self diagnosis and triage: audit study. *BMJ* 2015 Jul 08;351:h3480 [[FREE Full text](#)] [doi: [10.1136/bmj.h3480](https://doi.org/10.1136/bmj.h3480)] [Medline: [26157077](https://pubmed.ncbi.nlm.nih.gov/26157077/)]
7. Yu SW, Ma A, Tsang VH, Chung LS, Leung SC, Leung LP. Triage accuracy of online symptom checkers for Accident and Emergency Department patients. *Hong Kong J Emerg Med* 2020 Jul 1;27(4):217-222. [doi: [10.1177/1024907919842486](https://doi.org/10.1177/1024907919842486)]
8. Hill MG, Sim M, Mills B. The quality of diagnosis and triage advice provided by free online symptom checkers and apps in Australia. *Med J Aust* 2020 Jun;212(11):514-519. [doi: [10.5694/mja2.50600](https://doi.org/10.5694/mja2.50600)] [Medline: [32391611](https://pubmed.ncbi.nlm.nih.gov/32391611/)]
9. Ceney A, Tolond S, Glowinski A, Marks B, Swift S, Palser T. Accuracy of online symptom checkers and the potential impact on service utilisation. *PLoS One* 2021 Jul 16;16(7):e0254088 [[FREE Full text](#)] [doi: [10.1371/journal.pone.0254088](https://doi.org/10.1371/journal.pone.0254088)] [Medline: [34265845](https://pubmed.ncbi.nlm.nih.gov/34265845/)]
10. Campbell JL, Fletcher E, Britten N, Green C, Holt T, Lattimer V, et al. The clinical effectiveness and cost-effectiveness of telephone triage for managing same-day consultation requests in general practice: a cluster randomised controlled trial comparing general practitioner-led and nurse-led management systems with usual care (the ESTEEM trial). *Health Technol Assess* 2015 Feb;19(13):1-212 [[FREE Full text](#)] [doi: [10.3310/hta19130](https://doi.org/10.3310/hta19130)] [Medline: [25690266](https://pubmed.ncbi.nlm.nih.gov/25690266/)]
11. Donovan E, Wilcox CR, Patel S, Hay AD, Little P, Willcox ML. Digital interventions for parents of acutely ill children and their treatment-seeking behaviour: a systematic review. *Br J Gen Pract* 2020 Mar;70(692):e172-e178 [[FREE Full text](#)] [doi: [10.3399/bjgp20X708209](https://doi.org/10.3399/bjgp20X708209)] [Medline: [32041764](https://pubmed.ncbi.nlm.nih.gov/32041764/)]
12. Anhang Price R, Fagbuyi D, Harris R, Hanfling D, Place F, Taylor TB, et al. Feasibility of Web-based self-triage by parents of children with influenza-like illness: a cautionary tale. *JAMA Pediatr* 2013 Feb;167(2):112-118. [doi: [10.1001/jamapediatrics.2013.1573](https://doi.org/10.1001/jamapediatrics.2013.1573)] [Medline: [23254373](https://pubmed.ncbi.nlm.nih.gov/23254373/)]
13. Wallace W, Chan C, Chidambaram S, Hanna L, Iqbal FM, Acharya A, et al. The diagnostic and triage accuracy of digital and online symptom checker tools: a systematic review. *medRxiv (forthcoming)* 2021 Dec 21. [doi: [10.1101/2021.12.21.21268167](https://doi.org/10.1101/2021.12.21.21268167)]
14. Gottlieb K, Petersson G. Limited evidence of benefits of patient operated intelligent primary care triage tools: findings of a literature review. *BMJ Health Care Inform* 2020 May;27(1):e100114 [[FREE Full text](#)] [doi: [10.1136/bmjhci-2019-100114](https://doi.org/10.1136/bmjhci-2019-100114)] [Medline: [32385041](https://pubmed.ncbi.nlm.nih.gov/32385041/)]
15. Meyer AN, Giardina TD, Spitzmueller C, Shahid U, Scott TM, Singh H. Patient perspectives on the usefulness of an artificial intelligence-assisted symptom checker: cross-sectional survey study. *J Med Internet Res* 2020 Jan 30;22(1):e14679 [[FREE Full text](#)] [doi: [10.2196/14679](https://doi.org/10.2196/14679)] [Medline: [32012052](https://pubmed.ncbi.nlm.nih.gov/32012052/)]
16. Morse KE, Ostberg NP, Jones VG, Chan AS. Use characteristics and triage acuity of a digital symptom checker in a large integrated health system: population-based descriptive study. *J Med Internet Res* 2020 Nov 30;22(11):e20549 [[FREE Full text](#)] [doi: [10.2196/20549](https://doi.org/10.2196/20549)] [Medline: [33170799](https://pubmed.ncbi.nlm.nih.gov/33170799/)]
17. Verzantvoort NC, Teunis T, Verheij TJ, van der Velden AW. Self-triage for acute primary care via a smartphone application: practical, safe and efficient? *PLoS One* 2018 Jun 26;13(6):e0199284 [[FREE Full text](#)] [doi: [10.1371/journal.pone.0199284](https://doi.org/10.1371/journal.pone.0199284)] [Medline: [29944708](https://pubmed.ncbi.nlm.nih.gov/29944708/)]
18. Winn AN, Somai M, Fergestrom N, Crotty BH. Association of use of online symptom checkers with patients' plans for seeking care. *JAMA Netw Open* 2019 Dec 02;2(12):e1918561 [[FREE Full text](#)] [doi: [10.1001/jamanetworkopen.2019.18561](https://doi.org/10.1001/jamanetworkopen.2019.18561)] [Medline: [31880791](https://pubmed.ncbi.nlm.nih.gov/31880791/)]
19. Schmieding ML, Mörgeli R, Schmieding MA, Feufel MA, Balzer F. Benchmarking triage capability of symptom checkers against that of medical laypersons: survey study. *J Med Internet Res* 2021 Mar 10;23(3):e24475 [[FREE Full text](#)] [doi: [10.2196/24475](https://doi.org/10.2196/24475)] [Medline: [33688845](https://pubmed.ncbi.nlm.nih.gov/33688845/)]
20. Schmieding ML, Kopka M, Schmidt K, Schulz-Niethammer S, Balzer F, Feufel MA. Triage accuracy of symptom checker apps: five-year follow-up evaluation. *J Med Internet Res (forthcoming)* 2021. [doi: [10.2196/31810](https://doi.org/10.2196/31810)]
21. Schaefer KE, Chen JY, Szalma JL, Hancock PA. A meta-analysis of factors influencing the development of trust in automation: implications for understanding autonomy in future systems. *Hum Factors* 2016 May;58(3):377-400. [doi: [10.1177/0018720816634228](https://doi.org/10.1177/0018720816634228)] [Medline: [27005902](https://pubmed.ncbi.nlm.nih.gov/27005902/)]
22. Hoff KA, Bashir M. Trust in automation: integrating empirical evidence on factors that influence trust. *Hum Factors* 2015 May;57(3):407-434. [doi: [10.1177/0018720814547570](https://doi.org/10.1177/0018720814547570)] [Medline: [25875432](https://pubmed.ncbi.nlm.nih.gov/25875432/)]
23. de Visser EJ, Krueger F, McKnight P, Scheid S, Smith M, Chalk S, et al. The world is not enough: trust in Cognitive agents. *Proc Hum Factors Ergon Soc Annu Meet* 2012 Sep 01;56(1):263-267. [doi: [10.1177/1071181312561062](https://doi.org/10.1177/1071181312561062)]
24. de Visser EJ, Monfort SS, McKendrick R, Smith MA, McKnight PE, Krueger F, et al. Almost human: anthropomorphism increases trust resilience in cognitive agents. *J Exp Psychol Appl* 2016 Sep;22(3):331-349. [doi: [10.1037/xap0000092](https://doi.org/10.1037/xap0000092)] [Medline: [27505048](https://pubmed.ncbi.nlm.nih.gov/27505048/)]
25. Pak R, Fink N, Price M, Bass B, Sturre L. Decision support aids with anthropomorphic characteristics influence trust and performance in younger and older adults. *Ergonomics* 2012;55(9):1059-1072. [doi: [10.1080/00140139.2012.691554](https://doi.org/10.1080/00140139.2012.691554)] [Medline: [22799560](https://pubmed.ncbi.nlm.nih.gov/22799560/)]
26. Hertz N, Wiese E. Good advice is beyond all price, but what if it comes from a machine? *J Exp Psychol Appl* 2019 Sep;25(3):386-395. [doi: [10.1037/xap0000205](https://doi.org/10.1037/xap0000205)] [Medline: [30702316](https://pubmed.ncbi.nlm.nih.gov/30702316/)]

27. Baldauf M, Fröhlich P, Endl R. Trust Me, I'm a Doctor – user perceptions of AI-driven apps for mobile health diagnosis. In: Proceedings of the 19th International Conference on Mobile and Ubiquitous Multimedia. 2020 Nov Presented at: MUM '20; November 22-25, 2020; Essen, Germany p. 167-178. [doi: [10.1145/3428361.3428362](https://doi.org/10.1145/3428361.3428362)]
28. Symptomate - Check Your Symptoms Online. Symptomate. 2021. URL: <https://symptomate.com/> [accessed 2021-03-26]
29. Barragán-Montero A, Javaid U, Valdés G, Nguyen D, Desbordes P, Macq B, et al. Artificial intelligence and machine learning for medical imaging: a technology review. *Phys Med* 2021 Mar;83:242-256. [doi: [10.1016/j.ejmp.2021.04.016](https://doi.org/10.1016/j.ejmp.2021.04.016)] [Medline: [33979715](https://pubmed.ncbi.nlm.nih.gov/33979715/)]
30. Proposal for a regulation of the European parliament and of the council laying down harmonised rules on artificial intelligence (artificial intelligence act) and amending certain union legislative acts. European Commission. 2021 Apr 21. URL: <https://eur-lex.europa.eu/legal-content/EN/TXT/?qid=1623335154975&uri=CELEX%3A52021PC0206> [accessed 2021-12-01]
31. Lerch FJ, Prietula MJ, Kulik CT. The Turing effect: the nature of trust in expert systems advice. In: Feltovich PJ, Ford KM, Hoffman R, editors. *Expertise in Context: Human and Machine*. Cambridge, MA, USA: MIT Press; May 1997:417-448.
32. Parasuraman R, Sheridan TB, Wickens CD. Situation awareness, mental workload, and trust in automation: viable, empirically supported cognitive engineering constructs. *J Cogn Eng Decis Mak* 2008 Jun 01;2(2):140-160. [doi: [10.1518/155534308x284417](https://doi.org/10.1518/155534308x284417)]
33. Kopka M, Balzer F, Feufel MA, Schmieding ML. Triage Capability of Laypersons: Retrospective Exploratory Analysis. *JMIR Human Factors* (forthcoming) 2022 [[FREE Full text](#)]
34. Lee J, Moray N. Trust, control strategies and allocation of function in human-machine systems. *Ergonomics* 1992 Oct;35(10):1243-1270. [doi: [10.1080/00140139208967392](https://doi.org/10.1080/00140139208967392)] [Medline: [1516577](https://pubmed.ncbi.nlm.nih.gov/1516577/)]
35. Lee JD, Moray N. Trust, self-confidence, and operators' adaptation to automation. *Int J Hum Comput Stud* 1994 Jan;40(1):153-184. [doi: [10.1006/ijhc.1994.1007](https://doi.org/10.1006/ijhc.1994.1007)]
36. Yee N, Bailenson JN, Rickertsen K. A meta-analysis of the impact of the inclusion and realism of human-like faces on user experiences in interfaces. In: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems. 2007 Apr Presented at: CHI '07; April 28-May 3, 2007; San Jose, CA, USA p. 1-10. [doi: [10.1145/1240624.1240626](https://doi.org/10.1145/1240624.1240626)]
37. Online participant recruitment for surveys and market research. Prolific. 2021. URL: <https://www.prolific.co/> [accessed 2021-04-09]
38. Peer E, Brandimarte L, Samat S, Acquisti A. Beyond the Turk: alternative platforms for crowdsourcing behavioral research. *J Exp Soc Psychol* 2017 May;70:153-163. [doi: [10.1016/j.jesp.2017.01.006](https://doi.org/10.1016/j.jesp.2017.01.006)]
39. Casey LS, Chandler J, Levine AS, Proctor A, Strolovitch DZ. Intertemporal differences among MTurk workers: time-based sample variations and implications for online data collection. *SAGE Open* 2017 Jun 14;7(2):215824401771277. [doi: [10.1177/2158244017712774](https://doi.org/10.1177/2158244017712774)]
40. Ho CJ, Slivkins A, Suri S, Vaughan JW. Incentivizing high quality crowdwork. *ACM SIGecom Exch* 2015 Dec;14(2):26-34. [doi: [10.1145/2904104.2904108](https://doi.org/10.1145/2904104.2904108)]
41. Ada Health. Ada - Check Your Health. App Store. 2021. URL: <https://apps.apple.com/app/id1099986434> [accessed 2020-12-23]
42. Microsoft PowerPoint (Version 2019). Microsoft Corporation. 2021. URL: <https://www.microsoft.com/de-de/microsoft-365/powerpoint> [accessed 2021-04-04]
43. Serif Labs. Affinity Photo (Version 1.9.2). App Store. 2021. URL: <https://apps.apple.com/de/app/affinity-photo/id824183456?mt=12> [accessed 2021-04-04]
44. Linearity. Vectornator - Grafikdesign (Version 4.0.1). App Store. 2021. URL: <https://apps.apple.com/de/app/vectornator-grafikdesign/id1219074514> [accessed 2021-04-04]
45. Salanitri D, Hare C, Borsci S, Lawson G, Sharples S, Waterfield B. Relationship between trust and usability in virtual environments: an ongoing study. In: Proceedings of the 17th International Conference on Human-Computer Interaction: Design and Evaluation. 2015 Presented at: HCI '15; August 2-7, 2015; Los Angeles, CA, USA p. 49-59. [doi: [10.1007/978-3-319-20901-2_5](https://doi.org/10.1007/978-3-319-20901-2_5)]
46. Pak R, McLaughlin AC, Bass B. A multi-level analysis of the effects of age and gender stereotypes on trust in anthropomorphic technology by younger and older adults. *Ergonomics* 2014;57(9):1277-1289. [doi: [10.1080/00140139.2014.928750](https://doi.org/10.1080/00140139.2014.928750)] [Medline: [24935771](https://pubmed.ncbi.nlm.nih.gov/24935771/)]
47. Zou X, Lv C, Zhang J. The effect of group membership, system reliability and anthropomorphic appearance on user's trust in intelligent decision support system. In: Proceedings of the 17th International Conference on Human-Computer Interaction: Design and Evaluation. 2020 Jul Presented at: HCII '20; July 19-24, 2020; Copenhagen, Denmark p. 231-242. [doi: [10.1007/978-3-030-49183-3_18](https://doi.org/10.1007/978-3-030-49183-3_18)]
48. Questback. Umfragesoftware für Studierende und Wissenschaftler. Unipark. 2021. URL: <https://www.unipark.com/umfragesoftware/> [accessed 2021-04-04]
49. Jian JY, Bisantz AM, Drury CG. Foundations for an empirically determined scale of trust in automated systems. *Int J Cogn Ergon* 2000 Mar;4(1):53-71. [doi: [10.1207/S15327566IJCE0401_04](https://doi.org/10.1207/S15327566IJCE0401_04)]
50. Gutzwiller RS, Chiou EK, Craig SD, Lewis CM, Lematta GJ, Hsiung CP. Positive bias in the 'Trust in Automated Systems Survey'? An examination of the Jian et al. (2000) scale. *Proc Hum Factors Ergon Soc Annu Meet* 2019 Nov 20;63(1):217-221. [doi: [10.1177/1071181319631201](https://doi.org/10.1177/1071181319631201)]

51. van Dongen K, van Maanen PP. Under-reliance on the decision aid: a difference in calibration and attribution between self and aid. *Proc Hum Factors Ergon Soc Annu Meet* 2006 Oct 1;50(3):225-229. [doi: [10.1177/154193120605000304](https://doi.org/10.1177/154193120605000304)]
52. Muralidharan L, de Visser EJ, Parasuraman R. The effects of pitch contour and flanging on trust in speaking cognitive agents. In: *CHI '14 Extended Abstracts on Human Factors in Computing Systems*. 2014 Presented at: CHI EA '14; April 26-May 1, 2014; Toronto, Canada p. 2167-2172. [doi: [10.1145/2559206.2581231](https://doi.org/10.1145/2559206.2581231)]
53. Jessup SA. Measurement of the Propensity to Trust Automation. Wright State University. 2018. URL: https://corescholar.libraries.wright.edu/etd_all/2209/ [accessed 2022-04-25]
54. Norman CD, Skinner HA. eHEALS: the eHealth Literacy Scale. *J Med Internet Res* 2006 Nov 14;8(4):e27 [FREE Full text] [doi: [10.2196/jmir.8.4.e27](https://doi.org/10.2196/jmir.8.4.e27)] [Medline: [17213046](https://pubmed.ncbi.nlm.nih.gov/17213046/)]
55. R Core Team. R: a language and environment for statistical computing. R Foundation for Statistical Computing. 2021. URL: <https://www.R-project.org/> [accessed 2021-05-14]
56. Wickham H, Averick M, Bryan J, Chang W, McGowan LD, François R, et al. Welcome to the Tidyverse. *J Open Source Softw* 2019 Nov 21;4(43):1686. [doi: [10.21105/joss.01686](https://doi.org/10.21105/joss.01686)]
57. Lesnoff M, Lancelot R. aod: Analysis of Overdispersed Data. The Comprehensive R Archive Network. 2019. URL: <https://cran.r-project.org/web/packages/aod/index.html> [accessed 2021-05-28]
58. Champely S, Ekstrom C, Dalgaard P, Gill J, Weibelzahl S, Anandkumar A, et al. pwr: Basic Functions for Power Analysis. The Comprehensive R Archive Network. 2020. URL: <https://cran.r-project.org/package=pwr> [accessed 2021-05-28]
59. Using technology to ease the burden on primary care. Healthwatch Enfield. 2019 Jan. URL: https://www.healthwatch.co.uk/sites/healthwatch.co.uk/files/reports-library/20190122_Enfield_%20Using%20technology%20to%20ease%20the%20burden%20on%20primary%20care.pdf [accessed 2021-03-13]
60. Woodcock C, Mittelstadt B, Busbridge D, Blank G. The impact of explanations on layperson trust in artificial intelligence-driven symptom checker apps: experimental study. *J Med Internet Res* 2021 Nov 03;23(11):e29386 [FREE Full text] [doi: [10.2196/29386](https://doi.org/10.2196/29386)] [Medline: [34730544](https://pubmed.ncbi.nlm.nih.gov/34730544/)]
61. Miller S, Gilbert S, Virani V, Wicks P. Patients' utilization and perception of an artificial intelligence-based symptom assessment and advice technology in a British primary care waiting room: exploratory pilot study. *JMIR Hum Factors* 2020 Jul 10;7(3):e19713 [FREE Full text] [doi: [10.2196/19713](https://doi.org/10.2196/19713)] [Medline: [32540836](https://pubmed.ncbi.nlm.nih.gov/32540836/)]
62. Levine DM, Mehrotra A. Assessment of diagnosis and triage in validated case vignettes among nonphysicians before and after Internet search. *JAMA Netw Open* 2021 Mar 01;4(3):e213287 [FREE Full text] [doi: [10.1001/jamanetworkopen.2021.3287](https://doi.org/10.1001/jamanetworkopen.2021.3287)] [Medline: [33779741](https://pubmed.ncbi.nlm.nih.gov/33779741/)]
63. Jussupow E, Spohrer K, Heinzl A, Gawlitza J. Augmenting medical diagnosis decisions? An investigation into physicians' decision-making process with Artificial Intelligence. *Inf Syst Res* 2021 Sep;32(3):713-735. [doi: [10.1287/isre.2020.0980](https://doi.org/10.1287/isre.2020.0980)]
64. Cooper AA, Humphreys KR. The uncertainty is killing me: self-triage decision making and information availability. *E J Appl Psychol* 2008 Oct 14;4(1):1-6. [doi: [10.7790/ejap.v4i1.124](https://doi.org/10.7790/ejap.v4i1.124)]
65. Jessup SA, Schneider TR, Alarcon GM, Ryan TJ, Capiola A. The measurement of the propensity to trust automation. In: Chen JY, Fragomeni G, editors. *Virtual, Augmented and Mixed Reality: Applications and Case Studies*. Cham, Switzerland: Springer International Publishing; 2019:476-489.
66. Jungmann SM, Klan T, Kuhn S, Jungmann F. Accuracy of a Chatbot (Ada) in the diagnosis of mental disorders: comparative case study with lay and expert users. *JMIR Form Res* 2019 Oct 29;3(4):e13863 [FREE Full text] [doi: [10.2196/13863](https://doi.org/10.2196/13863)] [Medline: [31663858](https://pubmed.ncbi.nlm.nih.gov/31663858/)]
67. Ćirković A. Evaluation of four artificial intelligence-assisted self-diagnosis apps on three diagnoses: two-year follow-up study. *J Med Internet Res* 2020 Dec 04;22(12):e18097 [FREE Full text] [doi: [10.2196/18097](https://doi.org/10.2196/18097)] [Medline: [33275113](https://pubmed.ncbi.nlm.nih.gov/33275113/)]
68. Gilbert S, Fenech M, Upadhyay S, Wicks P, Novorol C. Quality of condition suggestions and urgency advice provided by the Ada symptom assessment app evaluated with vignettes optimised for Australia. *Aust J Prim Health* 2021 Oct;27(5):377-381. [doi: [10.1071/PY21032](https://doi.org/10.1071/PY21032)] [Medline: [34706813](https://pubmed.ncbi.nlm.nih.gov/34706813/)]
69. Semigran HL, Levine DM, Nundy S, Mehrotra A. Comparison of physician and computer diagnostic accuracy. *JAMA Intern Med* 2016 Dec 01;176(12):1860-1861. [doi: [10.1001/jamainternmed.2016.6001](https://doi.org/10.1001/jamainternmed.2016.6001)] [Medline: [27723877](https://pubmed.ncbi.nlm.nih.gov/27723877/)]

Abbreviations

- AI:** artificial intelligence
EFS: Enterprise Feedback Suite
OR: odds ratio
-

Edited by A Kushniruk; submitted 25.11.21; peer-reviewed by S Gilbert, G Petersson, C Lowe; comments to author 10.01.22; revised version received 09.02.22; accepted 06.03.22; published 03.05.22.

Please cite as:

Kopka M, Schmieding ML, Rieger T, Roesler E, Balzer F, Feufel MA

Determinants of Laypersons' Trust in Medical Decision Aids: Randomized Controlled Trial

JMIR Hum Factors 2022;9(2):e35219

URL: <https://humanfactors.jmir.org/2022/2/e35219>

doi: [10.2196/35219](https://doi.org/10.2196/35219)

PMID: [35503248](https://pubmed.ncbi.nlm.nih.gov/35503248/)

©Marvin Kopka, Malte L Schmieding, Tobias Rieger, Eileen Roesler, Felix Balzer, Markus A Feufel. Originally published in JMIR Human Factors (<https://humanfactors.jmir.org>), 03.05.2022. This is an open-access article distributed under the terms of the Creative Commons Attribution License (<https://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Human Factors, is properly cited. The complete bibliographic information, a link to the original publication on <https://humanfactors.jmir.org>, as well as this copyright and license information must be included.

Original Paper

A Mobile App for Prevention of Cardiovascular Disease and Type 2 Diabetes Mellitus: Development and Usability Study

Vera Helen Buss^{1,2}, PharmB, RPh, MRes; Marlien Varnfield¹, BSc, MSc, PhD; Mark Harris², MBBS, MD; Margo Barr², BSc, GCertTertTLP, MPH, PhD

¹Australian e-Health Research Centre, Commonwealth Scientific and Industrial Research Organisation, Herston, Australia

²Centre for Primary Health Care and Equity, University of New South Wales, Sydney, Australia

Corresponding Author:

Vera Helen Buss, PharmB, RPh, MRes
Australian e-Health Research Centre
Commonwealth Scientific and Industrial Research Organisation
Level 7, Surgical, Treatment and Rehabilitation Service
296 Herston Road
Herston, 4029
Australia
Phone: 61 732533603
Email: vera.buss@csiro.au

Abstract

Background: Cardiovascular disease (CVD) and type 2 diabetes mellitus (T2DM) are posing a huge burden on health care systems worldwide. Mobile apps can deliver behavior change interventions for chronic disease prevention on a large scale, but current evidence for their effectiveness is limited.

Objective: This paper reported on the development and user testing of a mobile app that aims at increasing risk awareness and engaging users in behavior change. It would form part of an intervention for primary prevention of CVD and T2DM.

Methods: The theoretical framework of the app design was based on the Behaviour Change Wheel, combined with the capability, opportunity, and motivation for behavior change system and the behavior change techniques from the Behavior Change Technique Taxonomy (version 1). In addition, evidence from scientific literature has guided the development process. The prototype was tested for user-friendliness via an iterative approach. We conducted semistructured interviews with individuals in the target populations, which included the System Usability Scale. We transcribed and analyzed the interviews using descriptive statistics for the System Usability Scale and thematic analysis to identify app features that improved utility and usability.

Results: The target population was Australians aged ≥ 45 years. The app included 4 core modules (risk score, goal setting, health measures, and education). In these modules, users learned about their risk for CVD and T2DM; set goals for smoking, alcohol consumption, diet, and physical activity; and tracked them. In total, we included 12 behavior change techniques. We conducted 2 rounds of usability testing, each involving 5 participants. The average age of the participants was 58 (SD 8) years. Totally, 60% (6/10) of the participants owned iPhone Operating System phones, and 40% (4/10) of them owned Android phones. In the first round, we identified a technical issue that prevented 30% (3/10) of the participants from completing the registration process. Among the 70% (7/10) of participants who were able to complete the registration process, 71% (5/7) rated the app above average, based on the System Usability Scale. During the interviews, we identified some issues related to functionality, content, and language and clarity. We used the participants' feedback to improve these aspects.

Conclusions: We developed the app using behavior change theory and scientific evidence. The user testing allowed us to identify and remove technical errors and integrate additional functions into the app, which the participants had requested. Next, we will evaluate the feasibility of the revised version of the app developed through this design process and usability testing.

(*JMIR Hum Factors* 2022;9(2):e35065) doi:[10.2196/35065](https://doi.org/10.2196/35065)

KEYWORDS

mobile health; behavior change intervention; primary prevention; health promotion; cardiovascular disease; diabetes mellitus, type 2; mobile phone

Introduction

Description of the Behaviors

Cardiovascular disease (CVD) and type 2 diabetes mellitus (T2DM) are 2 widely prevalent chronic conditions [1]. They are highly associated with unhealthy lifestyle, including tobacco smoking, alcohol consumption, poor diet, and physical inactivity. This means that preventative interventions that target these 4 behavioral risk factors can help reduce the risk of developing CVD and T2DM. According to the Global Burden of Disease Study 2019, these 4 factors are among the top behavioral risk factors for the total burden of disease [2]. In 2018, tobacco use contributed the most (8.6%) to the total disability-adjusted life years of Australians, followed by overweight and obesity (8.4%), dietary risk factors (5.4%), high blood pressure (5.1%), and alcohol use (4.5%) [3]. The authors of the report estimated that 38% of the burden of disease measured in the Australian Burden of Disease Study 2018 could have been prevented by reducing or avoiding exposure to the modifiable risk factors that were included in the study [3].

Overview of Existing Mobile Interventions for Chronic Disease Prevention

Mobile health interventions can be used to address these risk behaviors and help people reduce their risk for CVD and T2DM. Many apps have already been developed for the 4 risk behaviors. In a systematic review from 2020, Milne-Ives et al [4] concluded that there was no strong evidence to show that mobile apps can effectively improve health behaviors or outcomes, because only a few studies demonstrated statistically significantly better results in the intervention compared with the control group [4]. Similar results were reported by Palmer et al [5], Marcolino et al [6], Romeo et al [7], and Lunde et al [8]. We conducted a systematic review of mobile health-based multiple risk factor interventions for the prevention of CVD and T2DM [9]. The review included 3 studies on CVD prevention [10-12] and 6 on T2DM prevention [13-18]. Although the evidence was weak, the findings indicated that at least small to moderate reductions in CVD and T2DM risk can be achieved through mobile health interventions [9]. More recently, Redfern et al [19] published the results of a randomized controlled trial of an app-based intervention for CVD prevention. The intervention aimed at changes in diet, physical activity, smoking, and mental health, but not in alcohol consumption [19]. It showed borderline improvements in risk factors (blood pressure and lipids) and risk-related behaviors (physical activity and eHealth literacy) [19]. None of these interventions targeted CVD and T2DM prevention simultaneously, which is a gap that we aimed to address.

A recurring criticism by authors of systematic reviews is the low quality of evidence for the effectiveness of mobile health interventions [4-9,20-25]. Many have highlighted a lack of rigorous reporting on the theory underlying the intervention and the behavior change techniques included in the app [4,5,8,20-25]. Michie et al [26] explained that the interventions need to be accurately and fully described to subsequently understand which parts of behavior change interventions contributed to outcomes. Carraça et al [21] and Black et al [25]

found that effective behavior change techniques vary depending on the mode of delivery, for example, between face-to-face and digital interventions. This means that researchers cannot simply refer to the results of face-to-face behavior change interventions when deciding which behavior change techniques are to be included in the app design.

In addition, Palmer et al [5] found that many studies have focused on individual risk factors. In their systematic review of mobile-based interventions for the prevention of noncommunicable diseases, only 2 of the 71 studies targeted smoking, diet, and physical inactivity, and none targeted all the 4 behaviors (ie, smoking, diet, physical inactivity, and alcohol consumption) [5]. Noble et al [27] conducted a systematic review to investigate which behavioral risk factors were related. They identified that the 4 behaviors often occurred in clusters. Therefore, they argued in favor of multiple behavioral risk interventions [27]. Geller et al [28] highlighted that the behavioral risk factors for chronic conditions are overlapping, which means that successfully addressing these factors will reduce the risk for various chronic diseases.

Objectives

This study formed part of a larger project in which we aimed to develop and evaluate a mobile app-based intervention for CVD and T2DM risk awareness and prevention. The intervention's goal is to reduce the risk of both CVD and T2DM by targeting smoking, alcohol consumption, poor diet, and physical inactivity. In this paper, we have described the systematic development and refinement of the mobile app through usability testing.

Methods

Methodology

This study followed the methodology developed by Tombor et al [29] for the development of digital behavior change interventions. The methodology combines elements of the United Kingdom's Medical Research Council guidance [30], Multiphase Optimization Strategy [31], and Behaviour Change Wheel (BCW) [32]. Following these multiphase approaches, the development of the app-based intervention was divided into three phases—preparation, design, and piloting.

Preparation Phase

Step 1: Identify Target Behavior

This app-based intervention focused on 4 behaviors: smoking cessation, moderate alcohol intake, healthy diet (ie, adequate fruit and vegetable intake and avoidance of sugary drinks), and physical activity (ie, walking and other forms of exercise). These 4 behaviors are associated with decreased risk of CVD and T2DM [2]. Regarding smoking, the Australian National Strategy aims to prevent the uptake of smoking, encourage smokers to stop as soon as possible, and support former smokers to stay smoke-free [33]. Regarding alcohol, the Australian guidelines recommend ≤ 10 standard drinks per week and a maximum of 4 drinks on any day [34]. The Australian recommendations for a healthy diet include a combination of 5 serves of vegetables and 2 serves of fruits per day and no sugary drinks [35]. The

advice on exercise states at least 150 minutes of physical activity for people aged 45 to 64 years per week and at least 30 minutes on most, if not all, days for people aged ≥ 65 years [36]. We assumed that not everyone had to change all 4 risk behaviors. The specific target would depend on the individuals and their needs. Long-term behavior changes were required to lower CVD or T2DM risk [37-40].

Step 2: Identify the Theoretical Base

We used BCW as a theoretical base and combined it with the Behavior Change Technique Taxonomy (version 1) by Michie et al [41] to develop the app. Michie et al [32] developed BCW based on findings from a systematic review of existing frameworks of behavior change interventions. The researchers assessed the identified frameworks and addressed each of their limitations in a unified framework, namely the BCW [32]. Then, they tested the reliability with which the framework can be applied in practice [32]. BCW incorporates the capability, opportunity, and motivation for behavior change (COM-B) system [32], which describes the 3 components, capability, opportunity, and motivation, that jointly influence behavior. BCW consists of 3 layers that interact with each other [32]. The COM-B system builds the inner layer [32]. Capability describes physical and psychological factors that allow an individual to act on certain behavior [32]. Opportunity consists of physical and social factors that enable behavior [32]. Motivation comprises automatic or reflective thought processes that influence an individual's action [32]. The next layer consists of 9 intervention functions (education, persuasion, incentivization, coercion, training, enablement, modeling, environmental restructuring, and restrictions), followed by a layer of 9 policy categories (environmental and social planning, communication and marketing, legislation, service provision, regulation, fiscal measures, and guidelines) [32].

Step 3: Review Relevant Scientific Literature

We conducted a systematic literature review to assess the current evidence for the effectiveness of mobile health-based interventions in reducing the risk for CVD and T2DM, with a focus on multiple behavioral risk factor interventions [9]. In addition, we conducted a scoping review between August 2019 and August 2020 to identify relevant papers on behavior change, user engagement, and persuasion in the context of digital health, CVD and T2DM risk prediction, goal setting theory, and risk communication.

Step 4: Conduct Needs Assessment

The focus of the app-based intervention is the prevention of CVD and T2DM in Australian adults aged ≥ 45 years through behavior change. The Australian CVD risk guidelines defined CVD "collectively...as coronary heart disease, stroke and other vascular disease including peripheral arterial disease and renovascular disease" [42]. Diagnostic criteria depended on the specific condition. The Royal Australian Commission of General Practitioners defined T2DM as a "chronic and progressive medical condition that results from two major metabolic dysfunctions: insulin resistance and then pancreatic islet cell dysfunction causing a relative insulin deficiency" [43]. The diagnostic criteria comprised presentation of hyperglycemic

crisis, a single elevated fasting blood glucose level ≥ 7 mmol/L, a single hemoglobin A_{1c} $\geq 6.5\%$, or a random blood glucose level ≥ 11.1 mmol/L (the criteria is slightly different for asymptomatic individuals) [43]. We chose the age group for the intervention based on the advice of the Royal Australian Commission of General Practitioners for general practitioners (GPs) to conduct screening for risk factors and potentially initiate preventative measures in the healthy population ≥ 45 years [44]. For quantitative needs assessment, we reviewed the data from the Australian Burden of Disease Study 2015 [45] and the Australian National Health Survey 2017-18 [46]; these are summarized in Table S1 of [Multimedia Appendix 1](#) [21,23-25,45-55]. For qualitative needs assessment, we studied the audio and video presentations of people's real-life experiences of aging in Australia by Healthtalk Australia [56].

Design Phase

Step 5: Select Mode of Delivery

We will deliver the intervention via a mobile app. According to the National Health Survey 2017-18 [46], 89.7% of Australians aged ≥ 45 years stated that they owned a mobile phone or smartphone.

Step 6: Select Intervention Components

We selected BCW intervention functions following the affordability, practicability, effectiveness and cost-effectiveness, acceptability, side effects and safety, and equity (APEASE) criteria [57]. The criteria include whether the intervention is within an acceptable budget, whether it can be delivered as designed, whether it delivers desirable outcomes in practice, whether the benefit-cost ratio is favorable, whether relevant stakeholders consider it as appropriate, whether the risk-benefit ratio is favorable, and whether it narrows or widens disparities between different societal groups [57]. We selected 4 intervention functions: education (ie, "increasing knowledge or understanding"), persuasion (ie, "using communication to induce positive or negative feelings or stimulate action"), incentivization (ie, "creating an expectation of reward"), and enablement (ie, "increasing means/reducing barriers to increase capability [beyond education and training] or opportunity [beyond environmental restructuring]") from the BCW [57]. Furthermore, we picked 2 policy categories: communication and marketing, and service provision. We chose the intervention components based on previous experience with other apps and published literature.

Step 7: Specify the Intervention Content by Behavior Change Techniques

We used the Behavior Change Technique Taxonomy (version 1) by Michie et al [41] to select suitable behavior change techniques and connected them to the appropriate intervention functions, COM-B system components, and policy categories. We identified literature that reported on effective behavior change techniques in mobile health interventions. Subsequently, we developed an intervention strategy based on the selected behavior change techniques. Several systematic reviews and meta-analyses as well as other studies have been published, aiming to identify effective behavior change techniques for mobile health interventions, such as the studies by Carraça et

al [21], Qin et al [24], Van Rhoon et al [47], Kaner et al [23], Garnett et al [48], McCrabb et al [49], Black et al [25], Schroeé et al [54], and Asbjørnsen et al [55] (for an overview of the results, refer to Table S2 in [Multimedia Appendix 1](#)). This shows that, currently, there is no absolute answer as to which techniques are effective in practice; however, there is a clear indication for the effectiveness of self-regulatory strategies. Owing to this uncertainty, we could not simply draw on the results from such meta-analyses to select effective behavior change techniques for our intervention.

Step 8: Translate the Intervention Into App Features

In regular meetings, the research team and software engineers discussed the practical translation of the intervention into app features, focusing on user-friendliness and accessibility aspects.

Step 9: Design a Prototype App

We based the design of the prototype app on previous apps developed by members of the research team [58-60]. These apps have been validated by different stakeholders, including patients and clinicians. With a focus on the APEASE criteria [57], we set the goal of keeping the app design simple and user-friendly and using less internet data volume and less storage space on the smartphone. In addition, the software engineers developed the app such that it was compatible with iPhone Operating System (iOS) and Android systems. The prototype included 4 core modules: risk score, goal setting, health measures, and education.

Pilot Phase—Step 10: Conduct User Testing

In the next step, we tested the usability of the app iteratively. We anticipated requiring 2 to 3 cycles to remove all the major design issues. Each cycle consisted of 5 participants from the target population (aged ≥ 45 years, residing in Australia, fluent in written and spoken English, and owning a smartphone with internet access). We based the sample size calculation on previous studies [61-63]. We recruited participants through the institutional Twitter account and by contacting community groups (eg, community choirs, community gardens, and advocacy groups for older Australians). We offered participants a gift voucher worth Aus \$20 (US \$14) to thank them for their participation. After providing consent, participants received the study instructions, a link to download the app, a dummy profile, and the user guide ([Multimedia Appendix 2](#)) via email. The app was available in a test version only; therefore, the iOS users were required to download the TestFlight app first. Once the app was installed, we asked the participants to use the information provided in the dummy profile to register with the app. We did not collect any app data and asked participants to

use the dummy information because, at this stage, we were interested only in the user-friendliness of the app. We invited participants to explore the app further and to book a time for a feedback interview. A researcher (VHB) conducted the semistructured phone interviews (refer to [Multimedia Appendix 3](#) for the interview guide). The interviews consisted of questions about the downloading and registration processes and the System Usability Scale [64]. Then, VHB transcribed the interviews verbatim, analyzed the results of the System Usability Scale using descriptive statistics, and conducted a thematic analysis using NVivo (version 12; QSR International). As described by Neubeck et al [65], we classified the findings into three themes: functionality, content, or language and clarity. On the basis of the findings of the thematic analysis, we resolved the identified issues and added features to the app and user guide according to the participants' feedback. We repeated the steps until we achieved an adequate version of the app that we could use in a feasibility study.

Ethics Approval

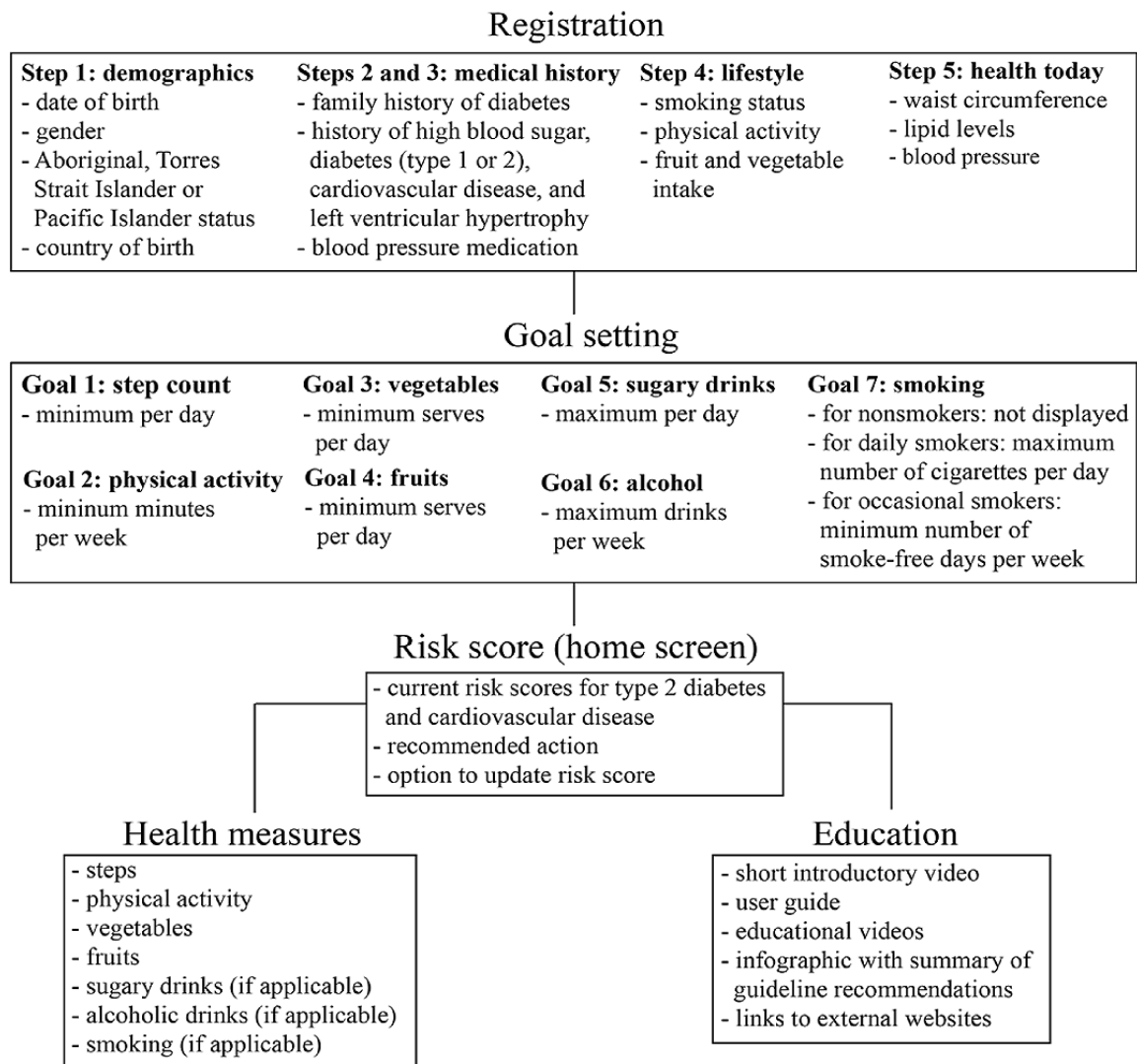
We received ethics approval from the University of New South Wales Australia Human Research Ethics Advisory Panel (approval number HC200069) and reciprocal approval from the Commonwealth Scientific and Industrial Research Organisation Health and Medical Human Research Ethics Committee (approval number 2020_041_RR).

Results

Registration Process and General App Features

The app included the following modules: registration, privacy policy and copyright, risk score, goal setting, health measures, and education. The first 2 modules were general app features, whereas the other 4 modules built the core intervention features. [Figure 1](#) shows the flow of the app, starting from registration. [Textbox 1](#) outlines the principles generated from the COM-B model, and [Table 1](#) shows the connection between the intervention modules and the selected intervention functions, the COM-B system components, and the intervention strategy for the app. It connects the 4 selected intervention functions (education, persuasion, incentivization, and enablement) with the specific behavior change techniques that we selected to achieve behavior change and the corresponding intervention strategies. [Figure 2](#) further outlines the connections between the selected intervention functions and the 4 app modules via the corresponding behavior change techniques. Refer to [Table S3](#) in [Multimedia Appendix 1](#) for more information on the design principles.

Figure 1. Flowchart of the app, starting from registration.



Textbox 1. Principles generated from the capability, opportunity, and motivation for behavior change model.

Psychological capability

- Educational videos and links to external websites with evidence-based health information to impart knowledge and train behavioral skills.
- Gradually making behavioral goals more difficult to train the behavioral skills.
- Providing values for goal setting and displaying last achieved values to train the cognitive skills.
- Advice on actions based on personal risks for cardiovascular disease (CVD) and type 2 diabetes mellitus (T2DM) as part of the visualization to improve understanding of adequate measures to achieve target behavior.

Physical capability

- Links to websites that provide healthy recipes and exercise instructions to develop skills in cooking and physical activity.

Social opportunity

- Advise for contacting general practitioner (included in risk score–related actions, educational videos, and advice if blood pressure or lipid levels not known) for social support.
- Links to websites including support programs (eg, for smoking cessation) for social support.

Physical opportunity

- Advise for contacting general practitioner (included in risk score–related actions, educational videos, and advice if blood pressure or lipid levels not known) who can check blood pressure levels and lipid levels or provide pharmacotherapy (eg, for hypertension or dyslipidemia) among others.

Automatic motivation

- Visualization of risk for CVD and T2DM displayed on the home screen of the app and the potential to change the risk based on health measures to elicit impulses and counterimpulses related to the target behavior.
- Announcing that the goals have been achieved in 3 consecutive weeks to trigger positive feelings about the behavioral goals.

Reflective motivation

- Facilitating self-monitoring of behavioral risk factors and reviewing the progress toward the self-set goals to increase understanding of own behavior and elicit positive or negative feelings about the behavioral goals.
- Highlighting the discrepancy between current behavior and goals to elicit positive or negative feelings and increase understanding of own capabilities about the behavioral goals.
- Encouraging self-reward after achieving weekly goals to elicit positive feelings about the behavioral goals.
- Providing feedback on personal risk of CVD and T2DM in the form of visualization to improve knowledge about own health and elicit positive feelings about the behavioral goals.
- Providing links to websites from credible sources and educational videos to increase knowledge and understanding about the target behavior.
- Setting goals to commit to target behavior and elicit positive feelings about it.

Table 1. Core intervention modules with corresponding intervention functions, COM-B^a system components, selected behavior change techniques, and intervention strategies.

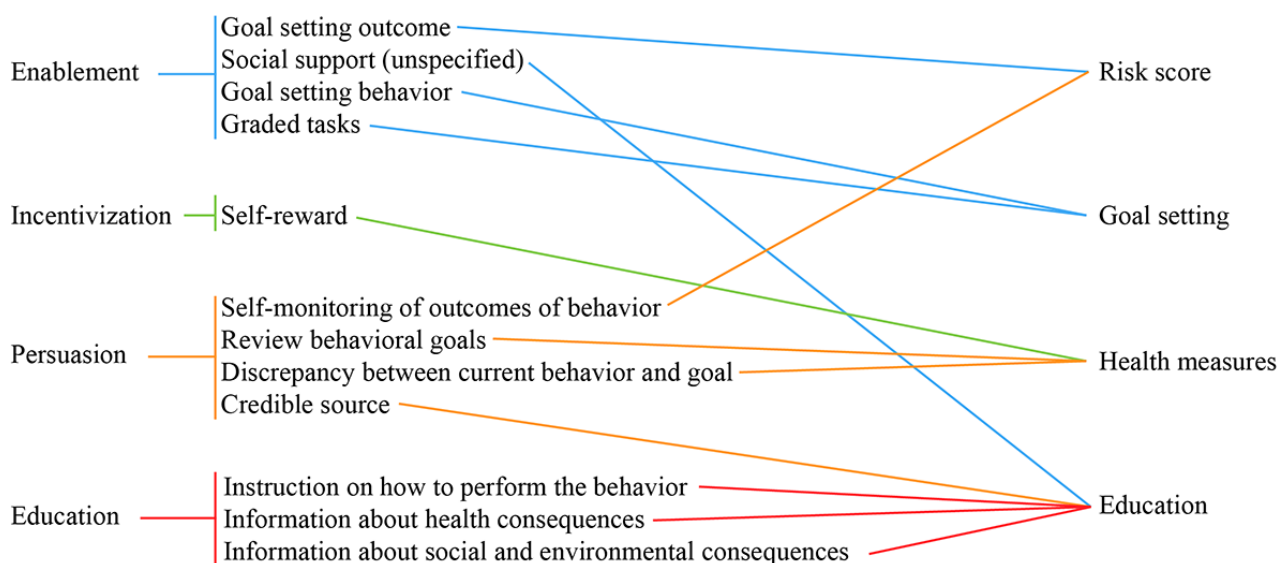
Module	Linked intervention functions	Key COM-B system components served by module	Selected behavior change techniques	Intervention strategies
Risk score	Enablement and persuasion	Automatic and reflective motivation	Self-monitoring of outcomes of behavior and goal setting (outcome)	Risk score visualization on home screen; outcome goal: low to moderate risk of CVD ^b and T2DM ^c ; and advice on actions based on personal risk, for example, advice to contact their general practitioner
Goal setting	Enablement	Psychological capability and reflective motivation	Goal setting (behavior) and graded tasks	Set behavioral goals for numbers of cigarettes, alcoholic drinks, fruit serves, vegetable serves, and sugary drinks, step count, and minutes of physical activity per day or week based on provided value ranges and advice to gradually make behavioral goals more difficult when they have been achieved in 3 consecutive weeks
Health measures	Persuasion and incentivization	Automatic and reflective motivation	Review behavioral goals, discrepancy between current behavior and goal, and self-reward	Icons and charts showing progress toward self-set daily or weekly goals, display discrepancies between current behavior and previously set goals through color-coding (red circle for negative counting or green circle for positive counting), display last achieved goals, encourage to reward themselves with an object or activity after they achieved their self-set goals, and congratulate when goals were achieved in 3 consecutive weeks
Education	Education, persuasion, and enablement	Physical and psychological capability	Instruction on how to perform the behavior, information about health consequences, information about social and environmental consequences, credible sources, and social support (unspecified)	Links to websites with health information and information about social, environmental, and emotional consequences; short videos on CVD, T2DM, and risk factors with advice to contact general practitioner; all information from credible sources (evidence-based); and links to support groups

^aCOM-B: capability, opportunity, and motivation for behavior change.

^bCVD: cardiovascular disease.

^cT2DM: type 2 diabetes mellitus.

Figure 2. Connections between the selected intervention functions and the 4 app modules via the corresponding behavior change techniques.



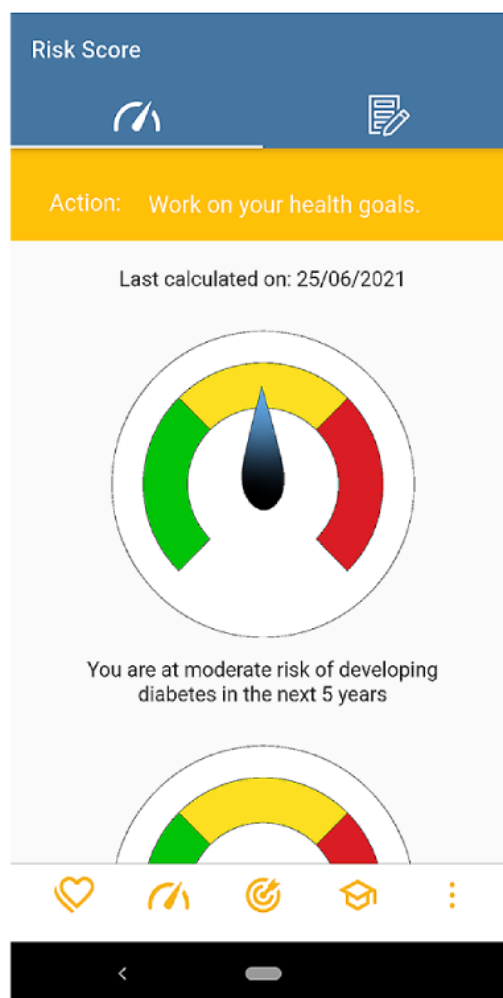
Risk Score

The long-term goal of the app-based intervention is primary prevention of CVD and T2DM. Therefore, risk presentations for both conditions built a central feature of the app (Figure 3). The risk scores that were embedded into the app were the 1991 version of the Framingham CVD risk score [66] and the Australian Type 2 Diabetes Risk Assessment Tool [67]. These are the standards currently used by clinicians in Australia and endorsed by the Royal Australian Commission of General Practitioners [43,44]. Users provided the information required for the risk calculation during the registration process. We designed the app such that the participants had the option to use Australian averages for lipids and blood pressure based on their age and sex if they did not know their values. Then, they received a recommendation to check the values with their GP. After completing the registration process, users saw their current risk of CVD and T2DM for the next 5 years.

Each time the users opened the app, they saw the risk score screen first. We hypothesized that this would create the impulse

to work on the behavioral goals to see a low or moderate disease risk displayed. We followed the principles that we identified during the risk communication scoping review [68-78]: keep the information simple and compact, use absolute instead of relative risk values, combine visuals with text, include information on action to take, and integrate a color scheme that is associated with different risk levels. In particular, the study by Reading Turchioe et al [71] influenced the final version of the visualization. The risk was stratified into 3 different levels (low, moderate, or high) and 3 corresponding actions (“Keep going. You are doing well”; “Work on your health goals”; or “Talk to your GP about your risk”). The recommended action focused on the higher of the 2 risks. Absolute risk values were not displayed because of the difficulty for users to interpret them. Users saw the date on which they had last updated their risk score and can update it at any time. Variables that do not change, such as gender and date of birth, were stored. Physical activity levels were collected through the health measures module. Users could update all other variables in the risk score module.

Figure 3. Screenshot of the risk score module.



Goal Setting

The app incorporated a goal setting function (Figure 4), which included proximal, specific goals related to the 4 behavioral

factors, smoking, alcohol intake, diet, and physical activity. We designed this module based on the following findings from our literature review, particularly, the principles from the Goal Setting Theory by Locke and Latham [79]. Miller et al [80]

stated that combining distal goals with proximal goals is particularly effective because proximal goals are perceived as an important step to achieving personally important distal goals. Locke and Latham [79] explained that specific goals increase self-efficacy and improve performance. Hence, the app incorporated a goal setting function that included proximal, specific goals related to the 4 behavioral factors. The motivation to change behavior is intended to arise from the goal to reduce the risks of CVD and T2DM, which is both a distal goal [79] and the behavior outcome of the intervention. Users had to self-set their behavioral goals, which, according to Locke and Latham [79], increases their self-efficacy compared with assigned goals. There is no goal related to body weight as the studies by Nothwehr and Yang [81] and Shilts et al [82] have shown that more specific goals such as diet and physical activity lead to better results than body weight. Regarding diet-related behavior change, Atkins and Michie [83] outlined that focusing on a specific diet-related target behavior was more effective

than focusing on the whole diet. In a study by Rohde et al [84], when asked to rate different food categories, participants voted fruits, vegetables, and sugary drinks as “easy-to-track.” Although whole grains play a significant role in a healthy diet, Foster et al [85] showed in an Australian survey that the general population had poor understanding of the subject. Hence, we assumed that comprehensive education on the subject would be required if it was to be incorporated into the app. The decision to track only sugary drink consumption and not discretionary food consumption was based on findings from 2 studies. Sui et al [86] reported that consumption of discretionary beverages has stronger correlation with high BMI in Australian adults than discretionary foods. Furthermore, Grieger et al [87] showed that raising fruit and vegetable intake can reduce discretionary food intake by subsidizing one for the other. Guided by these findings, we selected fruits, vegetables, and sugary drinks as the diet-related target behaviors.

Figure 4. Screenshot of the goal setting module.

The screenshot shows a mobile application interface for setting goals. At the top, there is a blue header with a close button (X), the title 'Goals', and a checkmark. Below the header, there are five goal categories, each with an icon and a dropdown menu for setting the goal frequency. The current values are all set to 0.

- Steps:** Icon of two feet. Dropdown: 'per day, at least *'. Latest: 0 steps.
- Physical Activity (mins):** Icon of a person walking. Dropdown: 'per week, at least *'. Latest: 0 min.
- Veggie Serves:** Icon of a carrot. Dropdown: 'per day, at least *'. Latest: 0 serves.
- Fruit Serves:** Icon of an apple. Dropdown: 'per day, at least *'. Latest: 0 serves.
- Sugary Drinks:** Icon of a glass with a straw. Dropdown: (empty). Latest: (empty).

At the bottom of the screen, there is a navigation bar with five icons: a heart, a leaf, a target, a graduation cap, and a vertical ellipsis. Below the navigation bar is a black bar with a back arrow and a home indicator.

To support users in their goal setting, the app displayed the last tracked values and offered a range of values for each goal to select from (eg, for vegetables, between 1-5 serves per day). The former helped people to set achievable goals, whereas the latter was consistent with the Australian guidelines. Further support for goal setting could be found in the educational

module. To ensure that every user set their goals at least once, they were directed to the goal setting module directly after completing the registration process. Regarding physical activity, users set 2 goals—step count and exercise. This provided users with both daily and weekly goals. Gouveia et al [88] have shown with physical activity that by using a default setting, people

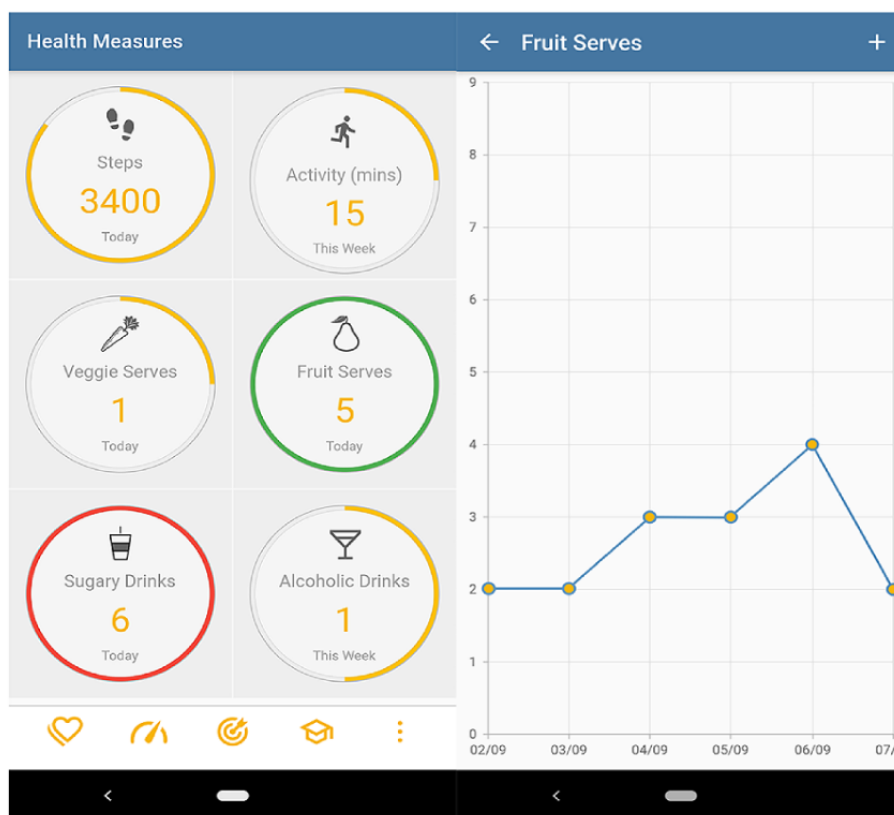
may keep this as their goal, even if they could achieve a more ambitious goal. In contrast, not everyone may be able to achieve the guideline recommendations, as argued by Kwasnicka et al [89]. Therefore, users could choose from a wide range of values that allowed for personalized and flexible goals, which Kwasnicka et al [89] argued to be important features in physical activity promotion. In the app, users could select that they did not drink any sugary or alcoholic drinks, which would automatically exclude these from the tracking function. Otherwise, they could select a value from the provided range. The smoking feature was personalized according to the smoking status that the users specified during the registration process.

Health Measures

In the health measures module (Figure 5), users could track their behavior and receive feedback on their progress toward the self-set goals. Locke and Latham [79] stated that the combination of single goal feedback with summary feedback has shown to be more effective than one of them alone. In the context of the app, the summary feedback (ie, feedback on the behavior outcome) was in the form of risk scores (ie, risk of CVD and T2DM), whereas the single goal feedback (ie, feedback on the four behaviors: smoking, alcohol, diet, and physical activity) was displayed in the health measures module. Each tracked goal was symbolized by an icon surrounded by a

circle. The circle showed the progress. With each step that brought the user closer to their goal, a part of the circle turned yellow. A fully colored circle indicated that the self-set goal had been achieved. For *positive* behaviors (eg, serves of fruits), the circle turned green, and for *negative* behaviors (eg, sugary drinks), the circle turned red. Two types of in-app push notifications encouraged self-reward and graded tasks (ie, increased the difficulty of the goal over time). First, when users achieved their weekly goals, they received a message encouraging them to reward themselves for their success (“You achieved your weekly goals, well done. Think of a way how you can reward yourself for that.”). The message addressed the incentivization intervention function and aimed to motivate users and increase their self-regulation, as postulated by Locke and Latham [79]. Second, when the users achieved the goals in 3 consecutive weeks, a message would pop up to encourage them to set a more ambitious goal (“You achieved your weekly goals 3 times in a row, well done. It’s time to set a more challenging goal.”). This addressed the enablement intervention function and was based on 2 principles. First, Locke and Latham [79] advocated for difficult but attainable goals. Second, Gordon et al [90] found that success in the first week of trying to achieve a new goal is highly determinant of the overall success toward this goal. Graphs showed users their behavior over time (Figure 5).

Figure 5. Screenshots of health measures module.



Education

The educational module (Figure 6) contained 5 videos explaining how different risk factors could increase an

individual’s risk of developing CVD and T2DM, the difference between modifiable and nonmodifiable risk factors, and behavioral and biomedical risk factors. Each video played for approximately 1 minute, was in plain language, and included

captions. Furthermore, for each behavior (smoking, alcohol, diet, and physical activity), there were links to external websites. These websites were Australian and contained evidence-based information about healthy lifestyle. The information was provided to help users to formulate strategies to attain their

goals, for example, quit lines for smoking cessation, recipes for a healthy diet, and exercise instructions for a home workout. An infographic (Figure 6) provided a quick overview of the guideline recommendations. This module also included the user guide and a short video that introduced the app.

Figure 6. Screenshots of the educational module.



Usability Testing Results

We conducted the usability tests between May 2021 and July 2021. In the second round of usability testing, we did not identify any major issues. Hence, we concluded the testing after the second round. A total of 12 individuals provided consent to participate in the study. Of these 12 individuals, 2 (17%) individuals were lost to follow-up and were not included in the analysis. The participants who completed the usability testing were, on average, aged 58 (SD 8) years, ranging from 47 to 67 years. Of the 10 participants, 2 (20%) participants identified as men and 8 (80%) identified as women. Of the 10 participants, 6 (60%) participants owned an iOS phone and 4 (40%) owned an Android phone. When asked about their app experience on a scale of 1 to 5 (1=very unexperienced and 5=very experienced), the average response was 4 (SD 0.9). In the first round, we identified a technical issue that prevented 30% (3/10) of the participants from completing the registration process. Table 2 presents the results of the System Usability Scale. Totally, 71% (5/7) of the participants who did not encounter the issue rated the usability of the app above average (ie, score >68 [64]). The median response for ease of use was strong agreement. The median response for confidence in using the app, frequent app use, and quickly learning to use the app was

agreement. The median response for inconsistencies in the app, cumbersome use, and the need to learn much before using the app was disagreement. The median response for the app being unnecessarily complex and requiring technical support for its use was strong disagreement. There were mixed results regarding how well the various functions in the app were integrated (median response was neutrality).

None of the participants faced issues while installing the app. All of them stated that the process was easy and quick. Those who did not experience the technical issue during the registration process said that the process was easy and quick. Participants said that it took them from <1 minute to couple of minutes to download the app and <5 minutes to register. During the interviews, the following issues regarding functionality were identified. A participant mentioned the need to adjust the font size for people with impaired vision. This functionality was already available, but we added an explanation to the user guide about how to adjust the font size through the phone settings. Another participant considered it “self-defeating” to have an extra file for the user guide. Hence, we integrated the user guide into the educational module of the app. In response to that participant’s comments, we included a short video in the educational module that explained the app’s features and their

purposes. Regarding language and clarity, we changed some icons and wording that the participants found unintuitive. For example, we changed the tick symbol to a return arrow and used the wording “tap to register” instead of “register.” Another finding regarding language and clarity was the ambiguity about the intended use of the app. A participant wanted the information directly in the app instead of the extra user guide. Another participant explained that it was not clear how often they were

supposed to enter their health measures. The participants in the user testing were only asked to download the app and complete the registration process and not to use it for a long period. Therefore, we had not provided them specific information about the frequency of use. For future users, we elaborated on the intended use of the app in the user guide and highlighted that we recommended regular use in the short introductory video.

Table 2. Results of the System Usability Scale (n=7).

Statement	Strongly agree, n (%)	Agree, n (%)	Neutral, n (%)	Disagree, n (%)	Strongly disagree, n (%)
I think that I would like to use this app frequently.	0 (0)	5 (71)	1 (14)	1 (14)	0 (0)
I found the app unnecessarily complex.	0 (0)	1 (14)	1 (14)	0 (0)	5 (71)
I thought the app was easy to use.	4 (57)	1 (14)	1 (14)	1 (14)	0 (0)
I think that I would need the support of a technical person to be able to use this app.	0 (0)	1 (14)	0 (0)	1 (14)	5 (71)
I found the various functions in this app were well integrated.	1 (14)	2 (29)	1 (14)	3 (43)	0 (0)
I thought there was too much inconsistency in this app.	0 (0)	1 (14)	1 (14)	3 (43)	2 (29)
I would imagine that most people would learn to use this app very quickly.	1 (14)	4 (57)	1 (14)	1 (14)	0 (0)
I found the app very cumbersome to use.	0 (0)	1 (14)	1 (14)	2 (29)	3 (43)
I felt very confident using the app.	3 (43)	3 (43)	0 (0)	1 (14)	0 (0)
I needed to learn a lot of things before I could get going with this app.	0 (0)	0 (0)	1 (14)	3 (43)	3 (43)

Regarding the content, a few participants expressed interest in the risk scores that were being calculated in the app. A participant stated the following:

And yet, it showed up that I was at moderate risk of getting diabetes in the next five years, and I thought: ‘What on Earth is that based on?’ and it undermined my confidence in the app. [P10]

To show that the risk scores were from credible sources, we included links to the websites for the risk scores in the educational module. Furthermore, some participants wanted to see time trends for the health measure. In response to their comments, we included a graphical display for each health measure that showed the progress over time. A participant expressed confusion about the external websites that we linked to in the education section. Another participant suggested a summary of the most important information about the health behaviors in the app in a visually appealing form. In response to these comments, we added some information about the external websites to the user guide and instructional video. In addition, we added an infographic to the educational module that summarized the guideline recommendations regarding the 4 behaviors. A few participants commented on the blood pressure and lipid levels that were required for the registration. Comments from the participants include the following:

I couldn’t answer my cholesterol levels. I couldn’t remember them.” [P6]

Why are they only asking about HDL? And why aren’t they asking about LDL? [P10]

Another participant explained the following regarding the registration process:

...my only thoughts when I was doing it that some people, uhm, wouldn’t know what a systolic or a diastolic pressure was and so that may be something that you may need. A little explanation of what that is. [P7]

We had already anticipated that users might not know their blood pressure or cholesterol levels. Hence, we included an option to use average values instead for the registration. The interviews further highlighted an issue that led us to add 2 educational videos explaining the relationship between disease risk, blood pressure, and cholesterol.

During the interviews, 3 of the participants mentioned commercial apps. Totally, 1 of them commented positively on the risk score module, but criticized that commercial apps would be better at incentivizing the user to continue by showing them how changes in their behavior influenced their risk and by providing incentives when the user achieved their target behavior. The remaining 2 participants said that many commercial apps automatically tracked many daily features, and concluded that this might not be required for our app. Comments from the participants include the following:

There’s lots of different apps in the market, isn’t there, to collect health data like that on a daily basis and monitor it. I don’t know if you need to go down that track of providing a trend, given that there’s so many other competitors in the market. [P10]

You know some of the fitness apps or whatever that I’ve been on they have actually almost too, too much stuff. I mean, I think this app is, is very good for kind of, you know, kind of like the basics. [P9]

Similarly, another participant explained the following:

I don't know what your demographic for the app is, but my parents are in their 90s and I think at least one of them would be able to, to use it, with a bit of help. [P1]

Discussion

Principal Findings

We developed an app that will form part of an intervention for the prevention of CVD and T2DM. The app's role in the intervention will be to make users aware of their disease risk and to engage them in healthy behavior. We developed the app around the principles of BCW to achieve a robust app construct. In total, we incorporated 12 behavior change techniques into the app to increase the capability, opportunity, and motivation of users to change their behavior. During the usability testing, participants ranked the usability of the app above average, based on the System Usability Scale. They stated that the app was easy and quick to download, basic in design, and easy to use. We used the participants' feedback to eliminate technical errors and adapt the app to their wishes and needs. Regarding the intervention, we do not anticipate that every user will adopt the *ideal* behaviors as described by the guidelines. For example, we do not expect that simply by using the app, a heavy drinker will stop consuming alcohol, a person who is obese will achieve normal weight, or a smoker will quit smoking. However, even small changes in behavior can decrease an individual's risk for CVD and T2DM. In addition, we do not consider the app as a stand-alone tool. Instead, we anticipate that app users will learn about their disease risk, risk behavior, and the connection between the 2 and that the app will help them seek information about where to receive help if needed (eg, from their GP or through support programs).

Comparison With Previous Studies

A similar study from Singapore that targeted coronary heart disease prevention via an app measured risk awareness, knowledge of risk factors, perceived stress levels, and heart-related lifestyle measures as outcomes [91]. Jiang et al [91] concluded that the intervention increased risk awareness and disease knowledge and the effects persisted for at least 6 months. They did not measure disease risk or incidence as outcomes [91]. The intervention of Jiang et al [91] differed from ours, as it focused on a 28-day time frame in which participants additionally received daily SMS text messages. In addition, the app included a stress management module, and the focus of the app was on written educational content including short quizzes [91,92]. The app features that we implemented were more diverse, including goal setting and tracking of different behavioral risk factors. Redfern et al [19] recruited Australians who are at moderate to high risk of CVD via GPs for their digital health intervention for CVD prevention. Similarities between their intervention and ours were that participants could learn about the relationship between CVD risk and their lifestyle, were encouraged to talk to their GP about these topics, and could set goals and monitor their lifestyle behavior [19]. Differences were that their intervention also focused on medication adherence and included data input from the

electronic health record [19]. Redfern et al [19] concluded that the intervention was not successful at improving medication adherence, which was the primary outcome.

Plotnikoff et al [93] developed an app-based intervention for T2DM prevention in Australia. In contrast to ours, a noteworthy proportion of their intervention was delivered in person in the form of cognitive training and exercise classes [93]. This limits the ability of the intervention to be scaled up. Block et al [14] developed a fully automated digital intervention for T2DM prevention in people who are prediabetic. The program consisted of weekly tailored goal setting and tracking of behaviors for the first 6 months and fortnightly for the following 6 months [14]. The intervention led to improved diabetes biomarkers and overall decreased T2DM risk [14]. In addition, it positively affected diet-related and physical activity-related behaviors [94]. The intervention differed from ours in that it included behavioral support for stress and sleep in addition to physical activity, diet, and weight loss [14]. It also comprised a website, automated phone calls, and emails [14]. Another difference was that it included social comparison features such as team competitions [14]. The theoretical base for competitions is the social upward comparison [95]. According to Spohrer et al [95], the social upward comparison theory is not compatible with the protection motivation theory because, in combination, they would lead to negative effects. We focused on aspects of the protection motivation theory. More specifically, the risk assessment module targets threat appraisal and the other modules target coping appraisal.

Implications and Future Research

We designed the app as simple as possible, so that it could be a tool for laypeople to use on their own. Ideally, it should encourage users to recognize their risks and make lifestyle-related changes without the direct need for medical or technical support. However, if they are at high risk of developing CVD or T2DM, engaging with the app should alert users and encourage them to seek help from their GP. We believe this is what sets our study apart from previous studies, which has tended to focus on people at high risk, includes the involvement of health care providers, or both. A recent systematic review of mobile health apps for the management of chronic conditions by Cucciniello et al [96] showed that the studies with additional human-led components did not have higher likelihood of positive effects on the outcomes for those in the intervention group. However, we believe that a few points should be considered when there is no direct involvement of health care professionals in the intervention. These include the appropriate promotion and uptake of the intervention to ensure that those who are likely to benefit are aware of and have access to it. In addition, the intervention should be designed such that the users who are at high risk will use it with medical supervision. Currently, we are in the process of evaluating the feasibility of this app-based intervention. Depending on the results of the feasibility study, we intend to conduct a study to evaluate the effectiveness of the app.

Strength and Limitations

A strength of our app was that the development process was guided by scientific evidence, with a focus on the APEASE

criteria. We provided a detailed description of the theoretical principles and design considerations, which showed transparency as opposed to the *It Seemed Like A Good Idea At The Time* principle [57]. This enabled the research team to understand which app features might be effective in changing the user's behavior. It also allows other researchers to replicate this study. Another strength was that we included feedback from potential users in the development process. A limitation was that the intervention focused on changes that were needed in the person rather than in the environment. The BCW from which we built the theoretical base comprised 9 intervention functions. We did not address 4 functions, such as coercion, restriction, environmental restructuring, and modeling, which, according to Michie et al [32], focus on external influences. Our app-based intervention focused on the personal agency of the participants. We limited diet-related risk factors to vegetables, fruits, and sugary drinks, which could potentially undermine the importance of other diet aspects, such as salt and whole grain intake. However, this was a conscious choice based on previous studies, which suggested that vegetables, fruits, and sugary drinks were easier to track than other diet-related behaviors. Another

limitation was that the app relies on user input. It does not automatically collect information, for example, step count or data from the electronic health record. We made these choices owing to reliability and privacy issues.

Conclusions

This paper describes the theoretical framework, design process, and usability testing of an app that will form the basis of an intervention for the primary prevention of CVD and T2DM. The app addressed the 3 behavior components, capability, opportunity, and motivation, which are core components of the BCW. In the usability testing, the participants rated the apps' usability as above average, according to the System Usability Scale. Most participants found the app easy to use, and they thought that most people would learn to use the app quickly. They also showed interest in using it frequently. After the user testing, some additional functions requested by the participants were integrated into the app. For example, a short introductory video and graphs that show the health measures over time were included. Next, we will use the revised version of the app that resulted from this design process and usability testing in a feasibility study.

Acknowledgments

This study was supported by a joint stipend from the University of New South Wales and the Commonwealth Scientific and Industrial Research Organisation. Furthermore, the authors would like to acknowledge the work of Karen Harrap (senior software engineer) and Derek Ireland (team leader) of the software engineering team, both at the Australian e-Health Research Centre, for providing their expertise and translating the theoretical framework and design concept into a viable product.

Authors' Contributions

VHB participated in the research design, conceptualization, and theoretical development of the app and writing of the manuscript. MV, MB, and MH contributed to the research design, conceptualization, and theoretical development of the app and revision of the manuscript. All the authors provided final approval for the version to be published.

Conflicts of Interest

None declared.

Multimedia Appendix 1

Additional information on the 4 behavioral risk factors, design principles, and evidence for the effectiveness of certain behavior change techniques.

[PDF File (Adobe PDF File), 103 KB - [humanfactors_v9i2e35065_app1.pdf](#)]

Multimedia Appendix 2

Android version of the user guide.

[PDF File (Adobe PDF File), 838 KB - [humanfactors_v9i2e35065_app2.pdf](#)]

Multimedia Appendix 3

Interview guide.

[PDF File (Adobe PDF File), 229 KB - [humanfactors_v9i2e35065_app3.pdf](#)]

References

1. GBD 2017 Disease and Injury Incidence and Prevalence Collaborators. Global, regional, and national incidence, prevalence, and years lived with disability for 354 diseases and injuries for 195 countries and territories, 1990-2017: a systematic analysis for the Global Burden of Disease Study 2017. *Lancet* 2018 Nov 10;392(10159):1789-1858 [FREE Full text] [doi: [10.1016/S0140-6736\(18\)32279-7](https://doi.org/10.1016/S0140-6736(18)32279-7)] [Medline: [30496104](https://pubmed.ncbi.nlm.nih.gov/30496104/)]

2. GBD 2019 Risk Factors Collaborators. Global burden of 87 risk factors in 204 countries and territories, 1990-2019: a systematic analysis for the Global Burden of Disease Study 2019. *Lancet* 2020 Oct 17;396(10258):1223-1249 [[FREE Full text](#)] [doi: [10.1016/S0140-6736\(20\)30752-2](https://doi.org/10.1016/S0140-6736(20)30752-2)] [Medline: [33069327](https://pubmed.ncbi.nlm.nih.gov/33069327/)]
3. Australian Burden of Disease Study 2018 – Key findings. Australian Institute of Health and Welfare. 2021 Aug 15. URL: <https://www.aihw.gov.au/reports/burden-of-disease/burden-of-disease-study-2018-key-findings/contents/about> [accessed 2021-11-15]
4. Milne-Ives M, Lam C, De Cock C, Van Velthoven MH, Meinert E. Mobile apps for health behavior change in physical activity, diet, drug and alcohol use, and mental health: systematic review. *JMIR Mhealth Uhealth* 2020 Mar 18;8(3):e17046 [[FREE Full text](#)] [doi: [10.2196/17046](https://doi.org/10.2196/17046)] [Medline: [32186518](https://pubmed.ncbi.nlm.nih.gov/32186518/)]
5. Palmer M, Sutherland J, Barnard S, Wynne A, Rezel E, Doel A, et al. The effectiveness of smoking cessation, physical activity/diet and alcohol reduction interventions delivered by mobile phones for the prevention of non-communicable diseases: a systematic review of randomised controlled trials. *PLoS One* 2018 Jan 5;13(1):e0189801 [[FREE Full text](#)] [doi: [10.1371/journal.pone.0189801](https://doi.org/10.1371/journal.pone.0189801)] [Medline: [29304148](https://pubmed.ncbi.nlm.nih.gov/29304148/)]
6. Marcolino MS, Oliveira JA, D'Agostino M, Ribeiro AL, Alkmim MB, Novillo-Ortiz D. The impact of mHealth interventions: systematic review of systematic reviews. *JMIR Mhealth Uhealth* 2018 Jan 17;6(1):e23 [[FREE Full text](#)] [doi: [10.2196/mhealth.8873](https://doi.org/10.2196/mhealth.8873)] [Medline: [29343463](https://pubmed.ncbi.nlm.nih.gov/29343463/)]
7. Romeo A, Edney S, Plotnikoff R, Curtis R, Ryan J, Sanders I, et al. Can smartphone apps increase physical activity? Systematic review and meta-analysis. *J Med Internet Res* 2019 Mar 19;21(3):e12053 [[FREE Full text](#)] [doi: [10.2196/12053](https://doi.org/10.2196/12053)] [Medline: [30888321](https://pubmed.ncbi.nlm.nih.gov/30888321/)]
8. Lunde P, Nilsson BB, Bergland A, Kværner KJ, Bye A. The effectiveness of smartphone apps for lifestyle improvement in noncommunicable diseases: systematic review and meta-analyses. *J Med Internet Res* 2018 May 04;20(5):e162 [[FREE Full text](#)] [doi: [10.2196/jmir.9751](https://doi.org/10.2196/jmir.9751)] [Medline: [29728346](https://pubmed.ncbi.nlm.nih.gov/29728346/)]
9. Buss VH, Leesong S, Barr M, Varnfield M, Harris M. Primary prevention of cardiovascular disease and type 2 diabetes mellitus using mobile health technology: systematic review of the literature. *J Med Internet Res* 2020 Oct 29;22(10):e21159 [[FREE Full text](#)] [doi: [10.2196/21159](https://doi.org/10.2196/21159)] [Medline: [33118936](https://pubmed.ncbi.nlm.nih.gov/33118936/)]
10. Rubinstein A, Miranda JJ, Beratarrechea A, Diez-Canseco F, Kanter R, Gutierrez L, GISMAL group. Effectiveness of an mHealth intervention to improve the cardiometabolic profile of people with prehypertension in low-resource urban settings in Latin America: a randomised controlled trial. *Lancet Diabetes Endocrinol* 2016 Jan;4(1):52-63. [doi: [10.1016/S2213-8587\(15\)00381-2](https://doi.org/10.1016/S2213-8587(15)00381-2)] [Medline: [26653067](https://pubmed.ncbi.nlm.nih.gov/26653067/)]
11. Gore MO, Krantz MJ, Albright K, Beaty B, Coronel-Mockler S, Bull S, et al. A controlled trial of mobile short message service among participants in a rural cardiovascular disease prevention program. *Prev Med Rep* 2019 Mar;13:126-131 [[FREE Full text](#)] [doi: [10.1016/j.pmedr.2018.11.021](https://doi.org/10.1016/j.pmedr.2018.11.021)] [Medline: [30568871](https://pubmed.ncbi.nlm.nih.gov/30568871/)]
12. Muntaner-Mas A, Vidal-Conti J, Borràs PA, Ortega FB, Palou P. Effects of a WhatsApp-delivered physical activity intervention to enhance health-related physical fitness components and cardiovascular disease risk factors in older adults. *J Sports Med Phys Fitness* 2017;57(1-2):90-102. [doi: [10.23736/S0022-4707.16.05918-1](https://doi.org/10.23736/S0022-4707.16.05918-1)] [Medline: [26364690](https://pubmed.ncbi.nlm.nih.gov/26364690/)]
13. Bender MS, Cooper BA, Flowers E, Ma R, Arai S. Filipinos Fit and Trim - a feasible and efficacious DPP-based intervention trial. *Contemp Clin Trials Commun* 2018 Dec;12:76-84 [[FREE Full text](#)] [doi: [10.1016/j.conctc.2018.09.004](https://doi.org/10.1016/j.conctc.2018.09.004)] [Medline: [30294699](https://pubmed.ncbi.nlm.nih.gov/30294699/)]
14. Block G, Azar KM, Romanelli RJ, Block TJ, Hopkins D, Carpenter HA, et al. Diabetes prevention and weight loss with a fully automated behavioral intervention by email, Web, and mobile phone: a randomized controlled trial among persons with prediabetes. *J Med Internet Res* 2015 Oct 23;17(10):e240 [[FREE Full text](#)] [doi: [10.2196/jmir.4897](https://doi.org/10.2196/jmir.4897)] [Medline: [26499966](https://pubmed.ncbi.nlm.nih.gov/26499966/)]
15. Fischer HH, Fischer IP, Pereira RI, Furniss AL, Rozwadowski JM, Moore SL, et al. Text message support for weight loss in patients with prediabetes: a randomized clinical trial. *Diabetes Care* 2016 Aug;39(8):1364-1370. [doi: [10.2337/dc15-2137](https://doi.org/10.2337/dc15-2137)] [Medline: [26861922](https://pubmed.ncbi.nlm.nih.gov/26861922/)]
16. Fukuoka Y, Gay CL, Joiner KL, Vittinghoff E. A novel diabetes prevention intervention using a mobile app: a randomized controlled trial with overweight adults at risk. *Am J Prev Med* 2015 Aug;49(2):223-237 [[FREE Full text](#)] [doi: [10.1016/j.amepre.2015.01.003](https://doi.org/10.1016/j.amepre.2015.01.003)] [Medline: [26033349](https://pubmed.ncbi.nlm.nih.gov/26033349/)]
17. Ramachandran A, Snehalatha C, Ram J, Selvam S, Simon M, Nanditha A, et al. Effectiveness of mobile phone messaging in prevention of type 2 diabetes by lifestyle modification in men in India: a prospective, parallel-group, randomised controlled trial. *Lancet Diabetes Endocrinol* 2013 Nov;1(3):191-198. [doi: [10.1016/S2213-8587\(13\)70067-6](https://doi.org/10.1016/S2213-8587(13)70067-6)] [Medline: [24622367](https://pubmed.ncbi.nlm.nih.gov/24622367/)]
18. Arens JH, Hauth W, Weissmann J. Novel app- and Web-supported diabetes prevention program to promote weight reduction, physical activity, and a healthier lifestyle: observation of the clinical application. *J Diabetes Sci Technol* 2018 Jul;12(4):831-838 [[FREE Full text](#)] [doi: [10.1177/1932296818768621](https://doi.org/10.1177/1932296818768621)] [Medline: [29584454](https://pubmed.ncbi.nlm.nih.gov/29584454/)]
19. Redfern J, Coorey G, Mulley J, Scaria A, Neubeck L, Hafiz N, et al. A digital health intervention for cardiovascular disease management in primary care (CONNECT) randomized controlled trial. *NPJ Digit Med* 2020 Sep 10;3:117 [[FREE Full text](#)] [doi: [10.1038/s41746-020-00325-z](https://doi.org/10.1038/s41746-020-00325-z)] [Medline: [32964140](https://pubmed.ncbi.nlm.nih.gov/32964140/)]

20. Duff OM, Walsh DM, Furlong BA, O'Connor NE, Moran KA, Woods CB. Behavior change techniques in physical activity eHealth interventions for people with cardiovascular disease: systematic review. *J Med Internet Res* 2017 Aug 02;19(8):e281 [FREE Full text] [doi: [10.2196/jmir.7782](https://doi.org/10.2196/jmir.7782)] [Medline: [28768610](https://pubmed.ncbi.nlm.nih.gov/28768610/)]
21. Carraça E, Encantado J, Battista F, Beaulieu K, Blundell J, Busetto L, et al. Effective behavior change techniques to promote physical activity in adults with overweight or obesity: a systematic review and meta-analysis. *Obes Rev* 2021 Jul;22 Suppl 4:e13258 [FREE Full text] [doi: [10.1111/obr.13258](https://doi.org/10.1111/obr.13258)] [Medline: [33949778](https://pubmed.ncbi.nlm.nih.gov/33949778/)]
22. Villinger K, Wahl DR, Boeing H, Schupp HT, Renner B. The effectiveness of app-based mobile interventions on nutrition behaviours and nutrition-related health outcomes: a systematic review and meta-analysis. *Obes Rev* 2019 Oct;20(10):1465-1484 [FREE Full text] [doi: [10.1111/obr.12903](https://doi.org/10.1111/obr.12903)] [Medline: [31353783](https://pubmed.ncbi.nlm.nih.gov/31353783/)]
23. Kaner EF, Beyer FR, Garnett C, Crane D, Brown J, Muirhead C, et al. Personalised digital interventions for reducing hazardous and harmful alcohol consumption in community-dwelling populations. *Cochrane Database Syst Rev* 2017 Sep 25;9:CD011479 [FREE Full text] [doi: [10.1002/14651858.CD011479.pub2](https://doi.org/10.1002/14651858.CD011479.pub2)] [Medline: [28944453](https://pubmed.ncbi.nlm.nih.gov/28944453/)]
24. Qin Y, Wang X, Namkoong K. A meta-analysis of the overall effect of mHealth physical activity interventions for weight loss and the moderating effect of behavioral change theories, techniques, and mobile technologies. *Mob Media Commun* 2021 Nov 16;10(2):337-359. [doi: [10.1177/20501579211054929](https://doi.org/10.1177/20501579211054929)]
25. Black N, Johnston M, Michie S, Hartmann-Boyce J, West R, Viechtbauer W, et al. Behaviour change techniques associated with smoking cessation in intervention and comparator groups of randomized controlled trials: a systematic review and meta-regression. *Addiction* 2020 Nov;115(11):2008-2020. [doi: [10.1111/add.15056](https://doi.org/10.1111/add.15056)] [Medline: [32196796](https://pubmed.ncbi.nlm.nih.gov/32196796/)]
26. Michie S, West R, Sheals K, Godinho CA. Evaluating the effectiveness of behavior change techniques in health-related behavior: a scoping review of methods used. *Transl Behav Med* 2018 Mar 01;8(2):212-224 [FREE Full text] [doi: [10.1093/tbm/ibx019](https://doi.org/10.1093/tbm/ibx019)] [Medline: [29381786](https://pubmed.ncbi.nlm.nih.gov/29381786/)]
27. Noble N, Paul C, Turon H, Oldmeadow C. Which modifiable health risk behaviours are related? A systematic review of the clustering of Smoking, Nutrition, Alcohol and Physical activity ('SNAP') health risk factors. *Prev Med* 2015 Dec;81:16-41. [doi: [10.1016/j.ypmed.2015.07.003](https://doi.org/10.1016/j.ypmed.2015.07.003)] [Medline: [26190368](https://pubmed.ncbi.nlm.nih.gov/26190368/)]
28. Geller K, Lippke S, Nigg CR. Future directions of multiple behavior change research. *J Behav Med* 2017 Feb;40(1):194-202. [doi: [10.1007/s10865-016-9809-8](https://doi.org/10.1007/s10865-016-9809-8)] [Medline: [27785652](https://pubmed.ncbi.nlm.nih.gov/27785652/)]
29. Tombor I, Shahab L, Brown J, Crane D, Michie S, West R. Development of SmokeFree Baby: a smoking cessation smartphone app for pregnant smokers. *Transl Behav Med* 2016 Dec;6(4):533-545 [FREE Full text] [doi: [10.1007/s13142-016-0438-0](https://doi.org/10.1007/s13142-016-0438-0)] [Medline: [27699682](https://pubmed.ncbi.nlm.nih.gov/27699682/)]
30. Craig P, Dieppe P, Macintyre S, Michie S, Nazareth I, Petticrew M, Medical Research Council Guidance. Developing and evaluating complex interventions: the new Medical Research Council guidance. *BMJ* 2008 Sep 29;337:a1655 [FREE Full text] [doi: [10.1136/bmj.a1655](https://doi.org/10.1136/bmj.a1655)] [Medline: [18824488](https://pubmed.ncbi.nlm.nih.gov/18824488/)]
31. Collins LM, Baker TB, Mermelstein RJ, Piper ME, Jorenby DE, Smith SS, et al. The multiphase optimization strategy for engineering effective tobacco use interventions. *Ann Behav Med* 2011 Apr;41(2):208-226. [doi: [10.1007/s12160-010-9253-x](https://doi.org/10.1007/s12160-010-9253-x)] [Medline: [21132416](https://pubmed.ncbi.nlm.nih.gov/21132416/)]
32. Michie S, van Stralen MM, West R. The behaviour change wheel: a new method for characterising and designing behaviour change interventions. *Implement Sci* 2011 Apr 23;6:42. [doi: [10.1186/1748-5908-6-42](https://doi.org/10.1186/1748-5908-6-42)] [Medline: [21513547](https://pubmed.ncbi.nlm.nih.gov/21513547/)]
33. Intergovernmental Committee on Drugs. National Tobacco Strategy 2012–2018: a strategy to improve the health of Australians by reducing smoking rates. Department of Health, Australian Government. 2012. URL: <https://www.health.gov.au/resources/publications/national-tobacco-strategy-2012-2018> [accessed 2021-11-15]
34. Australian guidelines to reduce health risks from drinking alcohol. National Health and Medical Research Council, Australian Government. 2020. URL: <https://www.nhmrc.gov.au/about-us/publications/australian-guidelines-reduce-health-risks-drinking-alcohol> [accessed 2021-11-15]
35. Australian dietary guidelines. Department of Health and Aging, National Health and Medical Research Council, Australian Government. 2013. URL: <https://www.health.gov.au/sites/default/files/australian-dietary-guidelines.pdf> [accessed 2022-02-14]
36. Physical activity and exercise guidelines for all Australians. Department of Health, Australian Government. 2021 May 07. URL: <https://tinyurl.com/2jywfrux> [accessed 2021-11-15]
37. Tuomilehto J, Lindström J, Eriksson JG, Valle TT, Hämäläinen H, Ilanne-Parikka P, Finnish Diabetes Prevention Study Group. Prevention of type 2 diabetes mellitus by changes in lifestyle among subjects with impaired glucose tolerance. *N Engl J Med* 2001 May 03;344(18):1343-1350. [doi: [10.1056/NEJM200105033441801](https://doi.org/10.1056/NEJM200105033441801)] [Medline: [11333990](https://pubmed.ncbi.nlm.nih.gov/11333990/)]
38. Knowler WC, Barrett-Connor E, Fowler SE, Hamman RF, Lachin JM, Walker EA, Diabetes Prevention Program Research Group. Reduction in the incidence of type 2 diabetes with lifestyle intervention or metformin. *N Engl J Med* 2002 Feb 07;346(6):393-403 [FREE Full text] [doi: [10.1056/NEJMoa012512](https://doi.org/10.1056/NEJMoa012512)] [Medline: [11832527](https://pubmed.ncbi.nlm.nih.gov/11832527/)]
39. Hooper L, Summerbell CD, Higgins JP, Thompson RL, Capps NE, Smith GD, et al. Dietary fat intake and prevention of cardiovascular disease: systematic review. *BMJ* 2001 Mar 31;322(7289):757-763 [FREE Full text] [doi: [10.1136/bmj.322.7289.757](https://doi.org/10.1136/bmj.322.7289.757)] [Medline: [11282859](https://pubmed.ncbi.nlm.nih.gov/11282859/)]
40. Simons-Morton DG, Calfas KJ, Oldenburg B, Burton NW. Effects of interventions in health care settings on physical activity or cardiorespiratory fitness. *Am J Prev Med* 1998 Nov;15(4):413-430. [doi: [10.1016/s0749-3797\(98\)00078-6](https://doi.org/10.1016/s0749-3797(98)00078-6)] [Medline: [9838981](https://pubmed.ncbi.nlm.nih.gov/9838981/)]

41. Michie S, Richardson M, Johnston M, Abraham C, Francis J, Hardeman W, et al. The behavior change technique taxonomy (v1) of 93 hierarchically clustered techniques: building an international consensus for the reporting of behavior change interventions. *Ann Behav Med* 2013 Aug;46(1):81-95. [doi: [10.1007/s12160-013-9486-6](https://doi.org/10.1007/s12160-013-9486-6)] [Medline: [23512568](https://pubmed.ncbi.nlm.nih.gov/23512568/)]
42. National Vascular Disease Prevention Alliance. Guidelines for the management of absolute cardiovascular disease risk 2012. Stroke Foundation. 2012. URL: <https://informme.org.au/Guidelines/Guidelines-for-the-assessment-and-management-of-absolute-CVD-risk> [accessed 2021-11-15]
43. General practice management of type 2 diabetes. The Royal Australian College of General Practitioners. 2016. URL: <https://static.diabetesaustralia.com.au/s/fileassets/diabetes-australia/5d3298b2-abf3-487e-9d5e-0558566fc242.pdf> [accessed 2021-11-15]
44. Guidelines for preventive activities in general practice. 9th edition. The Royal Australian College of General Practitioners. 2018. URL: <https://www.racgp.org.au/clinical-resources/clinical-guidelines/key-racgp-guidelines/view-all-racgp-guidelines/guidelines-for-preventive-activities-in-general-practice/preamble/introduction> [accessed 2021-11-15]
45. Australian Burden of Disease Study: impact and causes of illness and death in Australia 2015. Australian Institute of Health and Welfare. 2019 Jun 13. URL: <https://www.aihw.gov.au/reports-data/health-conditions-disability-deaths/burden-of-disease/data> [accessed 2021-11-15]
46. National Health Survey: first results, 2017-18. Australian Bureau of Statistics. 2019. URL: <https://www.abs.gov.au/AUSSTATS/abs@.nsf/DetailsPage/4364.0.55.0012017-18?OpenDocument> [accessed 2021-11-15]
47. Van Rhoon L, Byrne M, Morrissey E, Murphy J, McSharry J. A systematic review of the behaviour change techniques and digital features in technology-driven type 2 diabetes prevention interventions. *Digit Health* 2020 Mar 24;6:2055207620914427. [doi: [10.1177/2055207620914427](https://doi.org/10.1177/2055207620914427)] [Medline: [32269830](https://pubmed.ncbi.nlm.nih.gov/32269830/)]
48. Garnett C, Crane D, West R, Brown J, Michie S. Identification of behavior change techniques and engagement strategies to design a smartphone app to reduce alcohol consumption using a formal consensus method. *JMIR Mhealth Uhealth* 2015 Jun 29;3(2):e73 [FREE Full text] [doi: [10.2196/mhealth.3895](https://doi.org/10.2196/mhealth.3895)] [Medline: [26123578](https://pubmed.ncbi.nlm.nih.gov/26123578/)]
49. McCrabb S, Baker AL, Attia J, Skelton E, Twyman L, Palazzi K, et al. Internet-based programs incorporating behavior change techniques are associated with increased smoking cessation in the general population: a systematic review and meta-analysis. *Ann Behav Med* 2019 Feb 01;53(2):180-195. [doi: [10.1093/abm/kay026](https://doi.org/10.1093/abm/kay026)] [Medline: [29750240](https://pubmed.ncbi.nlm.nih.gov/29750240/)]
50. Arch A, Abou-Zahra S, Henry SL. Older Users Online: WAI Guidelines Address Older Users Web Experience. Web Accessibility Initiative. 2009. URL: <https://www.w3.org/WAI/posts/2009/older-users-online> [accessed 2022-02-14]
51. Senior Surfers: Diverse levels of digital literacy among older Australians. National Seniors Australia. 2019 Aug. URL: <https://nationalseniors.com.au/uploads/NationalSeniorsAustralia-SeniorSurfer-ResearchReport-2019.pdf> [accessed 2022-02-14]
52. Helbostad JL, Vereijken B, Becker C, Todd C, Taraldsen K, Pijnappels M, et al. Mobile health applications to promote active and healthy ageing. *Sensors (Basel)* 2017 Mar 18;17(3):622 [FREE Full text] [doi: [10.3390/s17030622](https://doi.org/10.3390/s17030622)] [Medline: [28335475](https://pubmed.ncbi.nlm.nih.gov/28335475/)]
53. Morey SA, Stuck RE, Chong AW, Barg-Walkow LH, Mitzner TL, Rogers WA. Mobile health apps: improving usability for older adult users. *Ergon Des* 2019 Apr 01;27(4):4-13. [doi: [10.1177/1064804619840731](https://doi.org/10.1177/1064804619840731)]
54. Schroë H, Van Dyck D, De Paepe A, Poppe L, Loh WW, Verloigne M, et al. Which behaviour change techniques are effective to promote physical activity and reduce sedentary behaviour in adults: a factorial randomized trial of an e- and m-health intervention. *Int J Behav Nutr Phys Act* 2020 Oct 07;17(1):127 [FREE Full text] [doi: [10.1186/s12966-020-01001-x](https://doi.org/10.1186/s12966-020-01001-x)] [Medline: [33028335](https://pubmed.ncbi.nlm.nih.gov/33028335/)]
55. Asbjørnsen RA, Smedsrød ML, Solberg Nes L, Wentzel J, Varsi C, Hjelmæsæth J, et al. Persuasive system design principles and behavior change techniques to stimulate motivation and adherence in electronic health interventions to support weight loss maintenance: scoping review. *J Med Internet Res* 2019 Jun 21;21(6):e14265 [FREE Full text] [doi: [10.2196/14265](https://doi.org/10.2196/14265)] [Medline: [31228174](https://pubmed.ncbi.nlm.nih.gov/31228174/)]
56. Ageing. Healthtalk Australia. 2021. URL: <https://www.healthtalkaustralia.org/experience-ageing/overview/> [accessed 2021-11-15]
57. Michie S, Atkins L, Gainforth HL. Changing behaviour to improve clinical practice and policy. In: Dias PC, Gonçalves A, Azevedo A, Lobo F, editors. *Novos desafios, novas competencias: contributos atuais da psicologia*. Braga, Portugal: Axioma - Publicações da Faculdade de Filosofia; 2016:41-60.
58. Varnfield M, Karunanithi M, Lee CK, Honeyman E, Arnold D, Ding H, et al. Smartphone-based home care model improved use of cardiac rehabilitation in postmyocardial infarction patients: results from a randomised controlled trial. *Heart* 2014 Nov;100(22):1770-1779 [FREE Full text] [doi: [10.1136/heartjnl-2014-305783](https://doi.org/10.1136/heartjnl-2014-305783)] [Medline: [24973083](https://pubmed.ncbi.nlm.nih.gov/24973083/)]
59. Varnfield M, Redd C, Stoney RM, Higgins L, Scolari N, Warwick R, et al. M♥THER, an mHealth system to support women with gestational diabetes mellitus: feasibility and acceptability study. *Diabetes Technol Ther* 2021 May;23(5):358-366 [FREE Full text] [doi: [10.1089/dia.2020.0509](https://doi.org/10.1089/dia.2020.0509)] [Medline: [33210954](https://pubmed.ncbi.nlm.nih.gov/33210954/)]
60. Gonzalez-Garcia M, Budd M, Varnfield M, Rogers R, Karunanithi M, Tan KS. PD-Buddy: mobile technology can provide support in homebased peritoneal dialysis to patients with chronic kidney disease. In: *Proceedings of the 25th Australian National Health Informatics Conference*. 2017 Presented at: HIC '17; August 6-9, 2017; Brisbane, Australia.

61. Molich R. A commentary of “How to specify the participant group size for usability studies: a practitioner’s guide” by R. Macefield. *J Usability Stud* 2010;5(3):124-128 [[FREE Full text](#)]
62. Stinson JN, Jibb LA, Nguyen C, Nathan PC, Maloney AM, Dupuis LL, et al. Development and testing of a multidimensional iPhone pain assessment application for adolescents with cancer. *J Med Internet Res* 2013 Mar 08;15(3):e51 [[FREE Full text](#)] [doi: [10.2196/jmir.2350](https://doi.org/10.2196/jmir.2350)] [Medline: [23475457](https://pubmed.ncbi.nlm.nih.gov/23475457/)]
63. Jibb LA, Cafazzo JA, Nathan PC, Seto E, Stevens BJ, Nguyen C, et al. Development of a mHealth real-time pain self-management app for adolescents with cancer: an iterative usability testing study [formula: see text]. *J Pediatr Oncol Nurs* 2017;34(4):283-294. [doi: [10.1177/1043454217697022](https://doi.org/10.1177/1043454217697022)] [Medline: [28376666](https://pubmed.ncbi.nlm.nih.gov/28376666/)]
64. System Usability Scale (SUS). Usability.: U.S. Department of Health & Human Services; 2021. URL: <https://www.usability.gov/how-to-and-tools/methods/system-usability-scale.html> [accessed 2021-11-15]
65. Neubeck L, Coorey G, Peiris D, Mulley J, Heeley E, Hersch F, et al. Development of an integrated e-health tool for people with, or at high risk of, cardiovascular disease: the Consumer Navigation of Electronic Cardiovascular Tools (CONNECT) web application. *Int J Med Inform* 2016 Dec;96:24-37. [doi: [10.1016/j.ijmedinf.2016.01.009](https://doi.org/10.1016/j.ijmedinf.2016.01.009)] [Medline: [26847070](https://pubmed.ncbi.nlm.nih.gov/26847070/)]
66. Anderson KM, Odell PM, Wilson PW, Kannel WB. Cardiovascular disease risk profiles. *Am Heart J* 1991 Jan;121(1 Pt 2):293-298. [doi: [10.1016/0002-8703\(91\)90861-b](https://doi.org/10.1016/0002-8703(91)90861-b)] [Medline: [1985385](https://pubmed.ncbi.nlm.nih.gov/1985385/)]
67. Chen L, Magliano DJ, Balkau B, Colagiuri S, Zimmet PZ, Tonkin AM, et al. AUSDRISK: an Australian type 2 diabetes risk assessment tool based on demographic, lifestyle and simple anthropometric measures. *Med J Aust* 2010 Feb 15;192(4):197-202. [doi: [10.5694/j.1326-5377.2010.tb03507.x](https://doi.org/10.5694/j.1326-5377.2010.tb03507.x)] [Medline: [20170456](https://pubmed.ncbi.nlm.nih.gov/20170456/)]
68. Lipkus IM, Hollands JG. The visual communication of risk. *J Natl Cancer Inst Monogr* 1999(25):149-163. [doi: [10.1093/oxfordjournals.jncimonographs.a024191](https://doi.org/10.1093/oxfordjournals.jncimonographs.a024191)] [Medline: [10854471](https://pubmed.ncbi.nlm.nih.gov/10854471/)]
69. Edwards A, Elwyn G, Mulley A. Explaining risks: turning numerical data into meaningful pictures. *BMJ* 2002 Apr 06;324(7341):827-830 [[FREE Full text](#)] [doi: [10.1136/bmj.324.7341.827](https://doi.org/10.1136/bmj.324.7341.827)] [Medline: [11934777](https://pubmed.ncbi.nlm.nih.gov/11934777/)]
70. Ancker JS, Senathirajah Y, Kukafka R, Starren JB. Design features of graphs in health risk communication: a systematic review. *J Am Med Inform Assoc* 2006;13(6):608-618 [[FREE Full text](#)] [doi: [10.1197/jamia.M2115](https://doi.org/10.1197/jamia.M2115)] [Medline: [16929039](https://pubmed.ncbi.nlm.nih.gov/16929039/)]
71. Reading Turchioe MR, Grossman LV, Myers AC, Baik D, Goyal P, Masterson Creber RM. Visual analogies, not graphs, increase patients' comprehension of changes in their health status. *J Am Med Inform Assoc* 2020 May 01;27(5):677-689 [[FREE Full text](#)] [doi: [10.1093/jamia/ocz217](https://doi.org/10.1093/jamia/ocz217)] [Medline: [31999316](https://pubmed.ncbi.nlm.nih.gov/31999316/)]
72. Lipkus IM. Numeric, verbal, and visual formats of conveying health risks: suggested best practices and future recommendations. *Med Decis Making* 2007;27(5):696-713. [doi: [10.1177/0272989X07307271](https://doi.org/10.1177/0272989X07307271)] [Medline: [17873259](https://pubmed.ncbi.nlm.nih.gov/17873259/)]
73. Peters E, Hibbard J, Slovic P, Dieckmann N. Numeracy skill and the communication, comprehension, and use of risk-benefit information. *Health Aff (Millwood)* 2007;26(3):741-748. [doi: [10.1377/hlthaff.26.3.741](https://doi.org/10.1377/hlthaff.26.3.741)] [Medline: [17485752](https://pubmed.ncbi.nlm.nih.gov/17485752/)]
74. Spiegelhalter D, Pearson M, Short I. Visualizing uncertainty about the future. *Science* 2011 Sep 09;333(6048):1393-1400. [doi: [10.1126/science.1191181](https://doi.org/10.1126/science.1191181)] [Medline: [21903802](https://pubmed.ncbi.nlm.nih.gov/21903802/)]
75. Garcia-Retamero R, Cokely ET. Designing visual aids that promote risk literacy: a systematic review of health research and evidence-based design heuristics. *Hum Factors* 2017 Jun;59(4):582-627. [doi: [10.1177/0018720817690634](https://doi.org/10.1177/0018720817690634)] [Medline: [28192674](https://pubmed.ncbi.nlm.nih.gov/28192674/)]
76. Damman OC, Vonk SI, van den Haak MJ, van Hooijdonk CM, Timmermans DR. The effects of infographics and several quantitative versus qualitative formats for cardiovascular disease risk, including heart age, on people's risk understanding. *Patient Educ Couns* 2018 Aug;101(8):1410-1418. [doi: [10.1016/j.pec.2018.03.015](https://doi.org/10.1016/j.pec.2018.03.015)] [Medline: [29559200](https://pubmed.ncbi.nlm.nih.gov/29559200/)]
77. Tao D, Yuan J, Qu X. Presenting self-monitoring test results for consumers: the effects of graphical formats and age. *J Am Med Inform Assoc* 2018 Aug 01;25(8):1036-1046 [[FREE Full text](#)] [doi: [10.1093/jamia/ocy046](https://doi.org/10.1093/jamia/ocy046)] [Medline: [29762686](https://pubmed.ncbi.nlm.nih.gov/29762686/)]
78. Gigerenzer G, Edwards A. Simple tools for understanding risks: from innumeracy to insight. *BMJ* 2003 Sep 27;327(7417):741-744 [[FREE Full text](#)] [doi: [10.1136/bmj.327.7417.741](https://doi.org/10.1136/bmj.327.7417.741)] [Medline: [14512488](https://pubmed.ncbi.nlm.nih.gov/14512488/)]
79. Locke EA, Latham GP. Building a practically useful theory of goal setting and task motivation. A 35-year odyssey. *Am Psychol* 2002 Sep;57(9):705-717. [doi: [10.1037//0003-066x.57.9.705](https://doi.org/10.1037//0003-066x.57.9.705)] [Medline: [12237980](https://pubmed.ncbi.nlm.nih.gov/12237980/)]
80. Miller RB, Brickman SJ. A model of future-oriented motivation and self-regulation. *Educ Psychol Rev* 2004 Mar;16(1):9-33. [doi: [10.1023/B:EDPR.0000012343.96370.39](https://doi.org/10.1023/B:EDPR.0000012343.96370.39)]
81. Nothwehr F, Yang J. Goal setting frequency and the use of behavioral strategies related to diet and physical activity. *Health Educ Res* 2007 Aug;22(4):532-538. [doi: [10.1093/her/cyl117](https://doi.org/10.1093/her/cyl117)] [Medline: [17032703](https://pubmed.ncbi.nlm.nih.gov/17032703/)]
82. Shilts MK, Horowitz M, Townsend MS. Goal setting as a strategy for dietary and physical activity behavior change: a review of the literature. *Am J Health Promot* 2004;19(2):81-93. [doi: [10.4278/0890-1171-19.2.81](https://doi.org/10.4278/0890-1171-19.2.81)] [Medline: [15559708](https://pubmed.ncbi.nlm.nih.gov/15559708/)]
83. Atkins L, Michie S. Designing interventions to change eating behaviours. *Proc Nutr Soc* 2015 May;74(2):164-170. [doi: [10.1017/S0029665115000075](https://doi.org/10.1017/S0029665115000075)] [Medline: [25998679](https://pubmed.ncbi.nlm.nih.gov/25998679/)]
84. Rohde A, Duensing A, Dawczynski C, Godemann J, Lorkowski S, Brombach C. An app to improve eating habits of adolescents and young adults (Challenge to Go): systematic development of a theory-based and target group-adapted mobile app intervention. *JMIR Mhealth Uhealth* 2019 Aug 12;7(8):e11575 [[FREE Full text](#)] [doi: [10.2196/11575](https://doi.org/10.2196/11575)] [Medline: [30903746](https://pubmed.ncbi.nlm.nih.gov/30903746/)]

85. Foster S, Beck E, Hughes J, Grafenauer S. Whole grains and consumer understanding: investigating consumers' identification, knowledge and attitudes to whole grains. *Nutrients* 2020 Jul 22;12(8):2170 [FREE Full text] [doi: [10.3390/nu12082170](https://doi.org/10.3390/nu12082170)] [Medline: [32708000](https://pubmed.ncbi.nlm.nih.gov/32708000/)]
86. Sui Z, Wong WK, Louie JC, Rangan A. Discretionary food and beverage consumption and its association with demographic characteristics, weight status, and fruit and vegetable intakes in Australian adults. *Public Health Nutr* 2017 Feb;20(2):274-281. [doi: [10.1017/S1368980016002305](https://doi.org/10.1017/S1368980016002305)] [Medline: [27572276](https://pubmed.ncbi.nlm.nih.gov/27572276/)]
87. Grieger JA, Wycherley TP, Johnson BJ, Golley RK. Discrete strategies to reduce intake of discretionary food choices: a scoping review. *Int J Behav Nutr Phys Act* 2016 May 06;13:57 [FREE Full text] [doi: [10.1186/s12966-016-0380-z](https://doi.org/10.1186/s12966-016-0380-z)] [Medline: [27151280](https://pubmed.ncbi.nlm.nih.gov/27151280/)]
88. Gouveia R, Karapanos E, Hassenzahl M. How do we engage with activity trackers? A longitudinal study of Habito. In: *Proceedings of the 2015 ACM International Joint Conference on Pervasive and Ubiquitous Computing*. 2015 Sep Presented at: UbiComp '15; September 7-11, 2015; Osaka, Japan p. 1305-1316. [doi: [10.1145/2750858.2804290](https://doi.org/10.1145/2750858.2804290)]
89. Kwasnicka D, Ntoumanis N, Sniehotta FF. Setting performance and learning goals is useful for active and inactive individuals, if goals are personalized and flexible: commentary on Swann et al. (2020). *Health Psychol Rev* 2021 Mar;15(1):51-55. [doi: [10.1080/17437199.2020.1762107](https://doi.org/10.1080/17437199.2020.1762107)] [Medline: [32608326](https://pubmed.ncbi.nlm.nih.gov/32608326/)]
90. Gordon M, Althoff T, Leskovec J. Goal-setting and achievement in activity tracking apps: a case study of MyFitnessPal. In: *Proceedings of the 2019 World Wide Web Conference*. 2019 May Presented at: WWW '19; May 13-17, 2019; San Francisco, CA, USA p. 571-582. [doi: [10.1145/3308558.3313432](https://doi.org/10.1145/3308558.3313432)]
91. Jiang Y, Jiao N, Nguyen HD, Lopez V, Wu VX, Kowitlawakul Y, et al. Effect of a mHealth programme on coronary heart disease prevention among working population in Singapore: a single group pretest-post-test design. *J Adv Nurs* 2019 Sep;75(9):1922-1932. [doi: [10.1111/jan.13980](https://doi.org/10.1111/jan.13980)] [Medline: [30786051](https://pubmed.ncbi.nlm.nih.gov/30786051/)]
92. Wang W, Zhang H, Lopez V, Wu VX, Poo DC, Kowitlawakul Y. Improving awareness, knowledge and heart-related lifestyle of coronary heart disease among working population through a mHealth programme: study protocol. *J Adv Nurs* 2015 Sep;71(9):2200-2207. [doi: [10.1111/jan.12668](https://doi.org/10.1111/jan.12668)] [Medline: [25879395](https://pubmed.ncbi.nlm.nih.gov/25879395/)]
93. Plotnikoff RC, Wilczynska M, Cohen KE, Smith JJ, Lubans DR. Integrating smartphone technology, social support and the outdoor physical environment to improve fitness among adults at risk of, or diagnosed with, type 2 diabetes: findings from the 'eCoFit' randomized controlled trial. *Prev Med* 2017 Dec;105:404-411. [doi: [10.1016/j.ypmed.2017.08.027](https://doi.org/10.1016/j.ypmed.2017.08.027)] [Medline: [28887192](https://pubmed.ncbi.nlm.nih.gov/28887192/)]
94. Block G, Azar KM, Romanelli RJ, Block TJ, Palaniappan LP, Dolginsky M, et al. Improving diet, activity and wellness in adults at risk of diabetes: randomized controlled trial. *Nutr Diabetes* 2016 Sep 19;6(9):e231 [FREE Full text] [doi: [10.1038/nutd.2016.42](https://doi.org/10.1038/nutd.2016.42)] [Medline: [27643726](https://pubmed.ncbi.nlm.nih.gov/27643726/)]
95. Spohrer K, Fallon M, Hoehle H, Heinzl A. Designing effective mobile health apps: does combining behavior change techniques really create synergies? *J Manag Inf Syst* 2021 Aug 06;38(2):517-545. [doi: [10.1080/07421222.2021.1912936](https://doi.org/10.1080/07421222.2021.1912936)]
96. Cucciniello M, Petracca F, Ciani O, Tarricone R. Development features and study characteristics of mobile health apps in the management of chronic conditions: a systematic review of randomised trials. *NPJ Digit Med* 2021 Oct 05;4(1):144 [FREE Full text] [doi: [10.1038/s41746-021-00517-1](https://doi.org/10.1038/s41746-021-00517-1)] [Medline: [34611287](https://pubmed.ncbi.nlm.nih.gov/34611287/)]

Abbreviations

APEASE: affordability, practicability, effectiveness and cost-effectiveness, acceptability, side effects and safety, and equity

BCW: Behaviour Change Wheel

COM-B: capability, opportunity, and motivation for behavior change

CVD: cardiovascular disease

GP: general practitioner

iOS: iPhone Operating System

T2DM: type 2 diabetes mellitus

Edited by A Kushniruk; submitted 19.11.21; peer-reviewed by J Meyer, J Ayre; comments to author 03.02.22; revised version received 26.02.22; accepted 26.03.22; published 10.05.22.

Please cite as:

Buss VH, Varnfield M, Harris M, Barr M

A Mobile App for Prevention of Cardiovascular Disease and Type 2 Diabetes Mellitus: Development and Usability Study

JMIR Hum Factors 2022;9(2):e35065

URL: <https://humanfactors.jmir.org/2022/2/e35065>

doi: [10.2196/35065](https://doi.org/10.2196/35065)

PMID: [35536603](https://pubmed.ncbi.nlm.nih.gov/35536603/)

©Vera Helen Buss, Marlien Varnfield, Mark Harris, Margo Barr. Originally published in JMIR Human Factors (<https://humanfactors.jmir.org>), 10.05.2022. This is an open-access article distributed under the terms of the Creative Commons Attribution License (<https://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Human Factors, is properly cited. The complete bibliographic information, a link to the original publication on <https://humanfactors.jmir.org>, as well as this copyright and license information must be included.

Original Paper

Audio Recording Patient-Nurse Verbal Communications in Home Health Care Settings: Pilot Feasibility and Usability Study

Maryam Zolnoori¹, MSc, PhD; Sasha Vergez², BSc, BA; Zoran Kostic³, PhD; Siddhartha Reddy Jonnalagadda⁴, PhD; Margaret V McDonald², MSc; Kathryn K H Bowles^{2,5}, PhD; Maxim Topaz^{1,2}, PhD

¹School of Nursing, Columbia University, New York, NY, United States

²Center for Home Care Policy & Research, Visiting Nurse Service of New York, New York, NY, United States

³Electrical Engineering, Columbia University, New York, NY, United States

⁴Amazon, Seattle, WA, United States

⁵School of Nursing, University of Pennsylvania, Philadelphia, NY, United States

Corresponding Author:

Maryam Zolnoori, MSc, PhD

School of Nursing

Columbia University

390 Fort Washington Avenue

New York, NY, 10033

United States

Phone: 1 317 515 1950

Email: mz2825@cumc.columbia.edu

Abstract

Background: Patients' spontaneous speech can act as a biomarker for identifying pathological entities, such as mental illness. Despite this potential, audio recording patients' spontaneous speech is not part of clinical workflows, and health care organizations often do not have dedicated policies regarding the audio recording of clinical encounters. No previous studies have investigated the best practical approach for integrating audio recording of patient-clinician encounters into clinical workflows, particularly in the home health care (HHC) setting.

Objective: This study aimed to evaluate the functionality and usability of several audio-recording devices for the audio recording of patient-nurse verbal communications in the HHC settings and elicit HHC stakeholder (patients and nurses) perspectives about the facilitators of and barriers to integrating audio recordings into clinical workflows.

Methods: This study was conducted at a large urban HHC agency located in New York, United States. We evaluated the usability and functionality of 7 audio-recording devices in a laboratory (controlled) setting. A total of 3 devices—Saramonic Blink500, Sony ICD-TX6, and Black Vox 365—were further evaluated in a clinical setting (patients' homes) by HHC nurses who completed the System Usability Scale questionnaire and participated in a short, structured interview to elicit feedback about each device. We also evaluated the accuracy of the automatic transcription of audio-recorded encounters for the 3 devices using the Amazon Web Service Transcribe. Word error rate was used to measure the accuracy of automated speech transcription. To understand the facilitators of and barriers to integrating audio recording of encounters into clinical workflows, we conducted semistructured interviews with 3 HHC nurses and 10 HHC patients. Thematic analysis was used to analyze the transcribed interviews.

Results: Saramonic Blink500 received the best overall evaluation score. The System Usability Scale score and word error rate for Saramonic Blink500 were 65% and 26%, respectively, and nurses found it easier to approach patients using this device than with the other 2 devices. Overall, patients found the process of audio recording to be satisfactory and convenient, with minimal impact on their communication with nurses. Although, in general, nurses also found the process easy to learn and satisfactory, they suggested that the audio recording of HHC encounters can affect their communication patterns. In addition, nurses were not aware of the potential to use audio-recorded encounters to improve health care services. Nurses also indicated that they would need to involve their managers to determine how audio recordings could be integrated into their clinical workflows and for any ongoing use of audio recordings during patient care management.

Conclusions: This study established the feasibility of audio recording HHC patient-nurse encounters. Training HHC nurses about the importance of the audio-recording process and the support of clinical managers are essential factors for successful implementation.

KEYWORDS

patients; HHC; communications; nurse; audio recording; device

Introduction

Patients' spoken language provides a window into a wide range of pathological entities, including pulmonary hypertension [1], respiratory obstruction [2], neurological disorders [3], and mental illnesses [4], enabling spoken language to act as a biomarker for screening patients with these diseases and symptoms. Recently, emergent studies have used established procedures in phonetics, speech sciences, and natural language processing to estimate changes in the phonatory and articulation of the patient's voice and to analyze semantic and pragmatic levels of language organization; these studies developed diagnostic and risk identification algorithms for the timely detection of diseases, particularly neurological and mental disorders [3,4].

Despite the promising findings of these studies, the audio recording of patients' spontaneous speech is not a part of routine clinical workflows. Most studies on patients' speech were cross-sectional and conducted in laboratory (controlled) settings, where patients were instructed to follow specific diagnostic and screening tests (eg, verbal fluency test and describing a positive or negative emotion) without having any interaction with clinicians [5-8]. These studies have several limitations, particularly in design and small sample sizes, which in turn limit the performance and generalizability of diagnosis or risk identification algorithms [3,4]. Integrating audio recordings of patient-clinician verbal communications during routine clinical encounters can potentially help resolve these limitations by creating an analytic pipeline of data sets to model subtle changes in patients' language, voice, emotion, interaction patterns, and engagement during clinical encounters. Such recordings can serve as a basis for developing intelligent clinical decision support systems that can help diagnose medical conditions or identify patients at risk for deterioration and negative outcomes.

Health care stakeholders' perspectives toward audio recording patient-clinician verbal communication have been discussed in previous studies. A recent study that investigated policies for audio recording patient-clinician encounters in the 49 largest health care systems in the United States found that despite physicians' willingness to audio record patient-clinician encounters, none of the health care systems had a dedicated policy or guidance for integrating audio recordings of patient-clinician encounters into clinical workflow [9]. In another study, Meeusen et al [10] explored patients' perspectives on the recording of their communication with neurosurgeons in an outpatient setting. Overall, patients had a positive perspective on recording, and they found it helpful. They also recommended that their future communication with clinicians be recorded [10]. In addition, Ball et al [11] also evaluated the perspectives of patients, clinicians, and clinic leaders toward audio recording patient-clinician encounters. The findings showed that patients saw audio recording as an opportunity to improve care. However, clinicians found it disruptive and

burdensome but appreciated its value for receiving low-stakes constructive feedback. Clinic leaders had a positive perspective on recording but were not prepared for its implementation in the clinical setting [11].

There is a growing consensus on the usability of home health care (HHC) technology as a significant factor affecting the use of technology in the HHC settings. HHC is a health care setting where services are provided by skilled practitioners (often registered nurses) to patients in their homes [12]. HHC patients are often clinically complex and vulnerable as they are generally older adults aged ≥ 65 years, with multiple chronic conditions such as Alzheimer disease and related disorders and respiratory and cardiac diseases. They are also at risk for negative outcomes such as emergency department visits and hospitalizations [13,14]. Audio recording and modeling of HHC patient-nurse encounters can allow the HHC team to enrich the documentation of patients' information in electronic health records (EHRs) and facilitate the development of high-performing clinical decision support systems to identify HHC patients at risk of health care deterioration (eg, Alzheimer disease), communication deficits (eg, aphasia), and negative outcomes (eg, emergency department visit). To assess the usability of technology such as audio recording of patient-nurse verbal communication in HHC settings, the International Organization for Standardization introduced 3 metrics: effectiveness, efficiency, and satisfaction [15]. Effectiveness mainly measures the extent to which the technology achieves its intended clinical goal, such as improvement in the performance of a diagnostic algorithm in diagnosing Alzheimer disease. Efficiency is related to the time and physical or mental effort needed to accomplish a task. Satisfaction measures the perceived usefulness and ease of use of technology from the perspective of health care stakeholders, including clinicians and patients. The Association for the Advancement of Medical Instrumentation and the Human Factors and Ergonomics Society emphasized the importance of HHC stakeholders' satisfaction and the conduction of a human factor analysis to improve usability [16]. HHC technology may be proven effective and efficient from the perspective of developers and researchers; however, HHC stakeholders, particularly patients and nurses, may find it unsatisfactory because of human factor issues such as substantial mental efforts or the time needed to learn about the technology.

Although some studies have reported on audio recording of clinical encounters mainly for patients' personal use (eg, recall of visit information) [17,18], few published insights are available on the usability of audio-recording devices and the accuracy of the recorded verbal communications. Most audio-recording devices can provide a sufficient level of quality for patients' personal use and documentation purposes; however, they do not offer the required accuracy for modeling the properties of a patient's verbal communication, particularly vocal (acoustic) parameters. In addition, the quality of recorded communication is highly dependent on the context of the clinical setting. In

clinical settings where patients and clinicians need to constantly move for physical examination or therapy, such as HHC settings [12], the location of the audio-recording device and the background noise can significantly affect the quality of the recorded communication and, in turn, the modeling verbal communication parameters. To integrate audio recordings into clinical workflows, it is also critical to consider patient and clinician attitudes and concerns regarding the audio-recording process. A patient's or clinician's negative attitude or discomfort during the recording process may disrupt the flow of treatment or result in the Hawthorne effect [19], affecting clinicians' communication patterns and treatment practices and preventing patients from sharing their actual concerns with clinicians.

To address the gaps in the literature, this study aimed to (1) evaluate the functionality and usability of several audio-recording devices for audio recording patient-nurse verbal communication in the HHC setting and (2) elicit the perspectives of HHC stakeholders (patients and nurses) about the facilitators

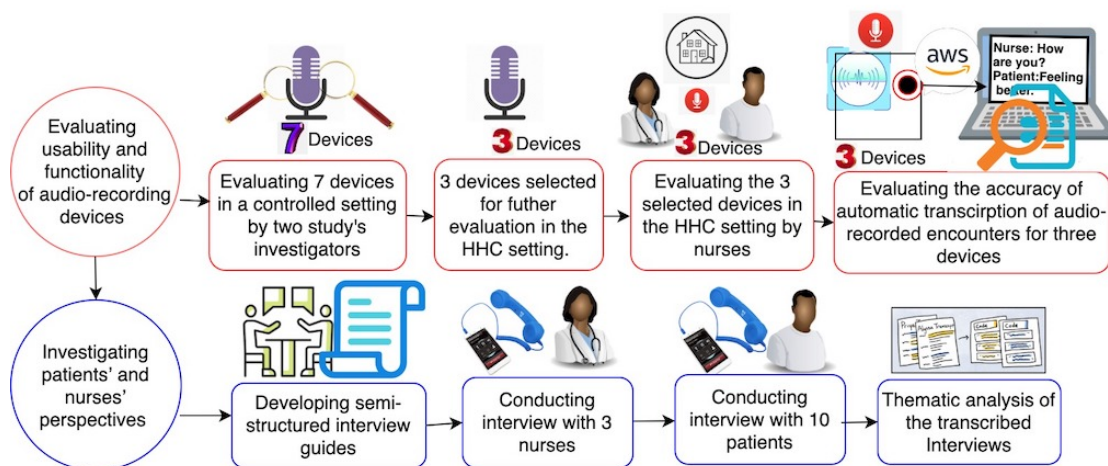
of and barriers to integrating audio recording into clinical workflows.

Methods

Study Setting

This descriptive feasibility study was conducted at the largest not-for-profit home health agency in the United States. The agency has approximately 10,600 staff, including 1470 nurses and >6500 home health aides. In 2019, the agency served >106,000 unique patients across >1.08 million clinical visits. A summary of patients' demographic information in 2019 showed that most of the Visiting Nurse Service of New York (VNSNY) patients were aged >65 years, predominantly women (63483/106,000, 59.89%), and almost half were African American or Hispanic (25,779/106,000, 24.32%, and 23,139/106,000, 21.83%, respectively). Figure 1 provides a schematic of the methodology used in this study.

Figure 1. A schematic view of the methodology of the study. HHC: home health care.



Ethics Approval

This study was approved by the Visiting Nurse Service of New York Institutional Review Board (reference #E20-003).

Evaluating the Functionality and Usability of Audio-Recording Devices

Overview

In the first phase of the study, we created a list of criteria for selecting the audio-recording devices. The criteria included the portability of the device, wearable features, functionality of the device (memory size and battery life), and voice activation

features. We reviewed the features of >50 audio-recording devices from different web-based sellers, such as Amazon, BestBuy, and SpyCenter. We selected 7 devices that met the criteria for the audio-recording device for the study. We evaluated the 7 devices for audio recording HHC patient-nurse verbal communication. All the devices were portable, with relatively simple operation, and could be used easily for recording verbal communication without disrupting the participants' movements during communication. The devices used were Black Vox 365, SOTA Surveillance-USR500, INSTAMIC PRO, Sony ICD-TX6, Mini Wristband Voice Activated Recorder, Apple Watch, and Saramonic Blink500 Pro B2. Figure 2 shows images of the audio-recording devices.

Figure 2. Audio-recording devices evaluated for audio recording of patient-nurse verbal communication.



Controlled Setting

In the first phase of evaluation, 2 study investigators (MZ and SV) used the 7 devices to audio record verbal communications between themselves in a controlled environment, which resembled the verbal communication between patients and nurses in HHC settings. The investigators evaluated the devices using 8 criteria: memory size; battery life; indicators for memory size, battery life, or recording status; automated voice activation feature; ease of device attachment to clothing; and accuracy of automatic transcription of recorded verbal communications as an indicator of voice quality (see the *Transcription Accuracy* section for evaluation metrics of automated voice activation features). The SOTA Surveillance-USR500, INSTAMIC PRO, Mini Wristband Voice Activated Recorder, and Apple Watch were excluded from further evaluation in the clinical settings because of their low battery life; low memory capacity; no indicators for recording (on or off), battery, and memory status; and low accuracy of automatic transcription of recorded communications. A total of 3 recorders, Black Vox 365 (further referred to as Vox), Sony ICD-TX6 (further referred to as Sony),

and Saramonic Blink500 Pro B2 Pro (further referred to as Saramonic), were included for further exploration in HHC because of the relatively high accuracy of automatic transcription, easy operation, good battery life, and memory capacity.

The Saramonic device comprises 2 wireless microphones that transmit the captured audio to a receiver connected to a recording device such as an iPod. The device has the feature of recording the communication between 2 people (eg, patient and nurse into the 2-channel recording), which can facilitate the process of separating the patient's voice from the nurse's voice for analysis purposes. By contrast, Sony and Vox embedded microphones with single-channel features. In contrast with Vox, Sony and Saramonic could be easily attached to the patients' and nurses' clothes and included indicators for recording status (on or off). Vox had a voice activation feature that was not available in Saramonic and Sony.

HHC Setting and Participant Recruitment

Next, we collaborated with 2 HHC nurses who used the 3 selected devices to audio record their verbal communication

with patients during HHC encounters. With HHC organizational support of the study and after institutional review board approval, we approached nurses through email and advertising at the VNSNY site. Interested nurses provided written consent. The consent form included information about the aim of the study and its potential risks and benefits. A research assistant (RA) trained the nurses to use the audio-recording devices.

Several different strategies were used to engage patient participants. The first strategy involved nurses providing flyers to patients with a brief description of the study. Nurses provided the RA with the name and contact information of patients who expressed interest. The second strategy was that after a waiver of authorization was granted, the RA reviewed the nurse's schedule in the EHR to identify patients on the nurse's caseload. The RA contacted the potential patient participants by phone and described the study's aim, potential risks, and benefits. For those who provided verbal consent, the RA mailed the consent form to the patients for their reference. Both nurse and patient participants received a gift card as a token of appreciation.

The participating nurses audio recorded their routine HHC encounters with consented patients. Audio-recorded encounters were uploaded to a secure server. Nurses also completed the System Usability Scale (SUS) questionnaires [20] for each device. The SUS is a robust and reliable instrument that measures a product's usability from the user's perspective. An example statement is, "I thought the system was easy to use."

The SUS provides scores from 0 (negative) to 100 (positive), with a standard average score of 68. A score >80 indicates excellent usability, whereas a score <68 indicates that the product has usability issues that are a cause for concern. To further evaluate the usability and functionality of the devices, we conducted semistructured, open-ended interviews with nurses to collect their opinions about the functionality and usability of the devices.

Transcription Accuracy

The quality of audio-recorded communications affects the transcription accuracy of a specific automatic speech recognition system such as the Amazon Web Service General Transcribe System (AWS-GTS) [21]. For example, the background noise and volume of the captured voice affect the quality of the transcription. We quantified the quality of audio recording patient-nurse verbal communications for the 3 devices using the automatic transcription and speaker identification features of the AWS-GTS. AWS-GTS was built on deep neural network models and trained on a large body of labeled (manually transcribed) verbal communication. We used this system as it is Health Insurance Portability and Accountability Act compliant and currently in use at the participating HHC organization.

We evaluated 2 components of transcription quality, word error rate (WER) and speaker identification accuracy, using the steps described in [Textbox 1](#).

Textbox 1. Transcription quality evaluation steps.

Step 1

For each device (Vox, Sony, and Saramonic), we randomly selected 3 audio-recorded home health care encounters.

Step 2

One of the investigators transcribed the audio files manually and assigned each utterance (defined as the uninterrupted part of the dialog expressed by one of the speakers) to an appropriate speaker (patient or nurse). Manual transcription was reviewed by a second investigator to ensure the quality of the transcription.

Step 3

All audio files were transcribed using the Amazon Web Service General Transcribe System application programming interface. The application programming interface returns the transcriptions as JSON files, including the start and end times of each transcribed word and the assigned speaker (eg, speaker 1 vs speaker 2) to each word. The transcribed words were joined to form an utterance using the type of speaker and converted to a Microsoft Excel sheet. We define an "utterance" as a continuous block of the uninterrupted speech of a single speaker.

Step 4

The quality of automatic transcriptions at the utterance level was compared with manual transcription and measured using the word error rate (WER). WER is a common metric for measuring the performance of speech recognition systems. It is computed based on the number of substitutions, insertions, and deletions that occur in a sequence of recognized words using a speech recognition system. The WER score starts from 0 (indicating no error in transcription) and can reach any score >1 depending on the length (number of words) in the utterance or document. For comparison, the average WER for human transcriptions is 0.04 (4%). Our earlier preliminary study for measuring the quality of transcription of an open-source automatic speech recognition system, Wav2Vec, on a subset of audio-recorded patient-nurse encounters provided a WER of 0.98 (98%). Wav2Vec is an unsupervised pretraining for speech recognition that learns representations of raw audio and was developed by the Facebook Company [22].

Second, we used the speaker identification feature of AWS-GTS to measure the accuracy of the automatic transcription of audio-recorded patient-nurse verbal communication by the devices. We expected that multiple-channel audio-recording devices (ie, Saramonic) would provide higher accuracy for speaker identification than that provided by single-channel devices (ie, Vox and Sony). To measure the accuracy of speaker identification, we used the following steps:

- Step 1: All manually transcribed audio files with assigned speakers at the utterance level (patient or nurse) were tokenized into words, and each word was linked to the assigned speaker.
- Step 2: The words of each manually transcribed utterance were mapped to the corresponding utterances and words provided by AWS-GTS.

- Step 3: We computed the percentage of words with accurate speaker identification with references to the total number of transcribed words in each audio file.

Investigating Patients' and Nurses' Perspectives About the Facilitators of and Barriers to the Integration of Audio-Recording of Verbal Communications Into HHC Clinical Workflows

Developing Semistructured Interview Guidelines

To understand the facilitators and barriers to the integration of audio recording of patient-nurse encounters into the clinical workflow, we conducted semistructured interviews with patients and nurses. The questions for nurses mainly covered their experience, concerns (eg, Hawthorne effect), potential benefits of recording (for both patients and clinicians), and their overall attitude toward the integration of audio-recording processing into HHC clinical workflows. The questions for patients covered their motivation to participate in the study, their concerns, and their attitudes toward audio recording their conversations with nurses. All the research questions were reviewed and discussed by the research team, 2 nurses with expertise in HHC services, a patient representative (familiar with HHC patients' characteristics, needs, preferences, and concerns), and 2 health informaticians. The study team ensured that the topics of the semistructured interviews would lead to the discovery of major facilitators and barriers to the development of a practical approach for audio recording patient-nurse encounters in a clinical setting. The questions for the semistructured interviews are presented in [Multimedia Appendix 1](#). Results of the interview guide development team discussions were summarized to generate an initial codebook for thematic analysis of the interviews.

Engagement of Study Participants in Qualitative Interviews

A total of 5 nurses audio recorded their communication with patients. However, during the study period, 40% (2/5) of nurses left the participating agency and were therefore not available to participate in the interviews. The nurses participated in audio-recorded encounters with 45 patients in this ongoing study. After securing additional consent, we conducted interviews with the remaining 60% (3/5) of nurses.

To reduce the likelihood of selection bias in creating a sample of HHC patients for the interviews, we used a stratified sampling technique. Using this technique, we stratified the pool of patients who participated in the audio recording of patient-nurse verbal communication (45 patients) based on the study nurse who recorded the encounter. Next, from each group of patients, the RA randomly contacted 2 patients for interviews. If the patient agreed to participate, the RA consented the patient for the study. This process was repeated until data saturation was achieved by interviewing a sample of 22% (10/45) of HHC patients.

To address the patient's potential concern about privacy and confidentiality of the information collected during the interviews and ensure that the patient expressed their unbiased perspectives, our RA consented the patients by explaining our commitment and strategies to protect the patients' privacy and confidentiality.

In addition, patients were informed that they could withdraw from the study (audio recording of patient-nurse verbal communication) without any consequences and had the freedom not to answer any questions.

In addition, our RA, who conducted the interviews, was well trained for effective communication (eg, active listening without interruption and speaking slowly and clearly) with VNSNY's patient population and had extensive experience in interviewing HHC patients for different qualitative studies. Gift cards were offered to both nurses and patients for participating in the qualitative interviews.

Thematic Analysis of the Interviews

The team created an initial codebook to summarize the open-ended interviews. Next, we used a thematic analysis approach for the systematic coding of the interviews. Thematic analysis is a qualitative descriptive approach for identifying, analyzing, and reporting themes within data [23-25]. The analysis phases were (1) familiarization with data, (2) generating initial codes, (3) data coding, (4) intercoder reliability, (5) searching for themes, (6) defining and naming themes, and (7) producing the report. To familiarize them with the data (step 1), 2 team investigators listened to and reviewed the transcribed interviews and used the initial codebook (step 2) to code transcriptions using a Microsoft Excel spreadsheet (step 3). The codes were assigned to informative passages of transcriptions (phrases, sentences, or paragraphs). Intercoding reliability was assured by dual coding of the first 2 interviews for each stakeholder (patient and nurse). Disagreements were resolved through discussions between the coders. Next, each investigator continued and coded the remaining interviews separately. When new codes emerged, they were discussed in our regular team meeting and added to the coding scheme if necessary (step 4). Final codes were collated into potential themes (broader and more abstract than codes) by undertaking an interpretative data analysis (step 5). The themes that emerged were further refined by analyzing the aspects of the data captured by each theme (step 6). This helped us to generate informative and clear names and definitions for a theme. The final report was produced by linking each theme to vivid and compelling interview quotes. Analysis of the patients' interviews and patient selection proceeded in an iterative process until data saturation was achieved (no new themes emerged from the data, and each theme was refined within a diverse sample).

Results

Demographic Information

Table 1 includes a summary of demographic information of the patients who participated in this study. The patients were predominantly men (6/10, 60%). Approximately 40% (4/10) of the study participants were Black, 30% (3/10) were White, and 20% (2/10) were Hispanic. One of the patients was not interested in declaring their race. More than half of the patients (6/10, 60%) were retired, and some (2/10, 20%) patients were disabled. All 3 nurses participating in this study were women: 2 (67%) were Black, and 1 (33%) was Hispanic. All 3 nurses had >5 years of clinical experience in HHC settings.

Table 1. Demographic information of patients (n=10) and nurses (n=3) participating in this study.

Demographics	Participants
Patients	
Gender (female), n (%)	4 (40)
Ethnicity, n (%)	
Black	4 (40)
White	3 (30)
Hispanic	2 (20)
Other	1 (10)
Age (years), mean (SD)	59.7 (16.25)
Employment, n (%)	
Employed	1 (10)
Unemployed	1 (10)
Retired	6 (60)
Disabled	2 (20)
Nurses	
Gender (female), n (%)	3 (100)
Ethnicity, n (%)	
Black	2 (67)
White	— ^a
Hispanic	1 (33)
Other	—
Work experience (years), n (%)	
<5	0 (0)
5-10	2 (67)
>10	1 (33)

^aNone of the nurses participated in this study was White.

Usability

Sony had the highest SUS score compared with Vox and Saramonic. Although the Saramonic device had a slightly lower SUS score than the SUS score of the Sony device, nurses found Saramonic easier in terms of approaching HHC patients for permission to audio record the verbal communication because of the appearance of the device and flexibility in attaching the microphone to the patient's clothes. In addition, the quality of audio-recorded communication using Saramonic was higher than that of the other 2 devices when measured using the WER

of automatic transcription and accuracy of speaker identification provided by AWS-GTS. As expected, the desirable usability feature of Vox was automatic voice-activated recording. However, this feature might compromise the patient's or nurse's privacy if the nurse forgets to pick up the device from the patient's home or if the nurse forgets to turn off the device, which in turn would start recording unrelated conversations. This feature was not available for Sony or Saramonic. Overall, we found that Saramonic is the most appropriate device for recording patient-nurse encounters with the highest SUS score and accuracy of speaker identification, as shown in [Table 2](#).

Table 2. Evaluation of 3 audio-recording devices in home health care settings by nurses and quality of audio-recorded files measured using Amazon Web Service General Transcribe System.

Device	System Usability Scale score	Overall opinion of the nurses about the device	Word error rate (%)	Accuracy of speaker identification (%)
Vox	42.5	The device lacked ease of usability because of the lack of indication of both the battery life and recording status.	38.4	67.3
Sony	78.75	The device was lightweight and easy to use; however, with the ongoing COVID-19 pandemic, the nurses were not comfortable putting the micro-phones near the face.	27	89.6
Saramonic	65	The device was simple to use. In addition, it was easier to approach patients using this device than with Sony.	26.3	91.3

Stakeholder Perspectives

We investigated the perspectives of HHC stakeholders (patients and nurses) toward integrating the audio-recording of patient-nurse verbal communications into the clinical workflow. Table 3 provides the thematic analysis of the interviews with patients. Overall, most patients were comfortable with the procedure of audio recording their communication with the

nurses. Some patients even mentioned that they completely forgot that their communication with the nurse was being recorded, and they did not have any concerns about sharing their concerns with nurses. In addition, most patients perceived the potential benefits of audio recording, particularly as a mechanism for recalling the nurses’ instructions, and they wanted the audio-recorded files to be shared with them.

Table 3. Summary of patients’ interviews.

Themes and subthemes	Common findings across patients	Differences across patients
Reasons for study enrollment	Most patients stated that their primary reasons for enrollment were to potentially help others and as they were satisfied with the services their nurses had provided them.	A patient had stated their primary reason for enrollment was the financial incentive.
Experience with recording		
Perceptions	Most patients had expressed feeling confident and that the device was not bothersome and was comfortable. Multiple patients stated that they forgot about the presence of the recorder soon after the visit started.	— ^a
Communication	All patients expressed that there was no effect on their communication with the nurses.	—

^aThere were no differences on the perceptions of nurses participated in this study.

Table 4 presents a summary of the thematic analysis of the nurses’ interviews. All nurses agreed that the recording device was convenient and easy to use, and they became used to the procedure after audio recording a few encounters (2-3 encounters each). However, some nurses disputed the idea of integrating the audio-recording of patient-nurse encounters into HHC clinical workflows because of the lengthening of the duration of the encounters, which cannot be easily suited to their current heavy workloads. For example, one of the nurses mentioned that “I don’t know. We are so busy. I remember there was a day I had 10 plus patients, that day I did not take the recorder. I don’t have time to do that.” When we investigated nurses’ opinions about the usefulness of audio-recorded verbal communication, some of them expressed doubt and argued that all the patient health-related information is documented in the

EHR system, and audio recording probably would not add more information. A nurse found the recording process helpful for managers in gaining insights into the quality of care in HHC settings. When asked about the impact of the audio recording on patient-nurse communication patterns, nurses stated that the recording would likely affect their communication patterns with patients. Specifically, audio recording may increase nurses’ willingness to establish more formal relationships with patients. Nurses also believed that patients’ communication patterns would change in the presence of audio-recording devices. When asked about the patients’ willingness to participate in this study, they responded that it depended on the patient’s personality, proactiveness in their care, and the complexity of their medical conditions.

Table 4. Summary of nurses' interviews.

Themes and subthemes	Common findings across nurses	Differences across nurses
Experience with recording		
Device usability	Most nurses had a good experience and stated that they began to become comfortable after a few uses (recording 1-2 encounters).	A nurse highlighted the difficulty in finding places to clip the microphone and attach the device in one home health care encounter.
Future use	When asked if this would become a standardized process, most nurses expressed dismay, citing that it would add more time to the visits with their already heavy workload.	— ^a
Potential use of recording for other purposes	Most nurses expressed doubts about its usefulness because of their present, heavy workload and no need to use it to look back at visits as everything is documented during the visit.	A nurse expressed the usefulness in a management aspect where it can potentially help with the assignment of the workload. It would help give insights into the visits that they would not otherwise see.
Patient encounters		
Communication	Most nurses said that the recording could affect the way patients communicate with nurses (Hawthorne effect). Most nurses said that the recording could affect the way the nurses develop relationships with their patients. In other words, it may increase the nurses' willingness to establish a more formal relationship with patients.	—
Participation	—	All the nurses answered differently. It could be dependent on the patients' personality, proactiveness in their care, or the type of case they are in (chronic care or wound care).

^aThere was no difference across nurses.

Discussion

Principal Findings

This study is the first to develop a practical approach for audio recording HHC patient-nurse verbal communication and evaluate the feasibility of integrating this approach into the clinical workflow. We showed that the type of recording device can have a differential effect on the accuracy of downstream tasks built on audio-recorded patient-nurse encounters, such as automatic transcription and provided speaker identification. In addition, our results suggest that HHC stakeholders' attitudes toward the recording process are a crucial factor affecting the successful integration of the audio recording of patient-nurse verbal communications into the clinical workflow.

Selecting an appropriate recording device is of great importance to designing and implementing an effective approach for audio recording clinical encounters in HHC settings. Some devices are lightweight, can be easily attached to the participants' clothes, and have a very simple operation mechanism with one on or off button. However, these devices often have functionality issues, such as short battery life or limited memory size, which reduce their usability by increasing the number of clinicians or administrative staff. The INSTAMIC PRO device is an example of such a device with high ease of use but low functionality because of its small memory capacity. In contrast, some devices are easy to use with good functionality; however, because of their size or weight, they cannot be easily attached to a patient's or clinician's clothes. Vox is an example of a device that needs to be set on a flat surface (eg, a table) during the process of recording. Although this solution is practical for recording

patient-clinician encounters where no movement is required during communication, it is not an appropriate solution in HHC settings where there is a constant required movement of the patient or nurse for physical examination or treatment. The patient's or nurse's movement will change their position with reference to the device location and, therefore, would affect the quality of audio-recorded communication, in turn affecting downstream tasks built on the audio data (eg, extracting some linguistic or acoustic features from the patient's speech).

In addition, the quality of the audio-recorded communication by devices substantially affects the quality of automatic transcription and speaker identification provided by an automatic speech recognition system. This is particularly important when the goal of the study is to develop an automatic analytic pipeline for processing and modeling patient-nurse verbal communication to develop a risk identification or diagnostic algorithm (eg, a diagnostic algorithm for Alzheimer disease). Regarding the possibility of background noise in HHC settings, which comes from different sources such as television, air conditioner, or a caregiver's speaking, it is important to set the microphone as close to the patient's and nurse's mouths to reduce the possibility of background noise captured by the device's microphone. An option for reducing background noise is to use a unidirectional (cardioid) overhead microphone; however, as this study was conducted during the COVID-19 pandemic and the study participants were recognized as being at high risk for COVID-19, we avoided any devices that touched the patient's face. Unidirectional microphones pick up audio (eg, the patient's voice) from only the front compared with omnidirectional microphones that pick up audio from all directions. Another reason that convinced us not to use overhead microphones was

the risk of patient discomfort during the recording procedure when the microphone touches the patient's face.

Another important feature of audio-recording devices is the number of channels with the ability to separate audio tracks for each individual participating in the communication. Devices with this type of feature usually have multiple microphones that are used by several individual speakers. The voice captured by each microphone is transmitted to an individual channel. This feature is important for downstream tasks, especially the differentiation of the patient's voice from the nurse's voice. Among all the devices evaluated in this study, the Saramonic device was the only device that included this feature; consequently, it had a higher score for speaker identification when it was measured using AWS-GTS. Overall, because of the better quality of automatic transcription and speaker identification (measured using AWS-GTS) of verbal communications recorded by this device and the high score for usability, we selected this device for audio recording further encounters beyond the pilot assessment.

Understanding HHC patients' and nurses' perspectives toward audio recording is a key determinant of the successful integration of the audio recording of patient-nurse verbal communication into the HHC clinical workflow. Overall, patients found the process of audio recording to be satisfactory and convenient, with minimal impact on their communication with nurses. Patients expressed that they were able to freely share their concerns and health care issues with nurses, and they even forgot about the presence of audio devices soon after the visit started. This expression implies that the Hawthorne effect on patients' communication patterns with nurses was minimal. It also shows the practicality of the recording procedure designed for this study.

The Hawthorne effect refers to a study participant's reactivity, in which the participant changes an aspect of their behavior in response to their awareness of being observed by the study's investigator [19]. At the beginning of the study, we briefly educated the patients about the audio-recording procedure and our strategy for protecting their privacy and confidentiality. We believe that this training is particularly essential for reducing the possibility of patient discomfort and, in turn, the Hawthorne effect.

Similar to patients, nurses found the audio-recording procedure satisfactory and easy to learn. However, in contrast to patients, nurses expressed that audio recording of HHC encounters can affect their practice and communication patterns in the HHC settings. Some nurses suggested that their communication with patients might become more *formal* rather than personal. Therefore, we may conclude that the presence of a recording device can introduce a Hawthorne effect on nurses' communication with patients. Nurses expressed concern about sharing audio-recorded encounters with supervisors for evaluation purposes. Owing to this concern, we experienced difficulties in recruiting nurses to audio record their communication with patients; however, after educating nurses about protecting their privacy and confidentiality, 5 nurses agreed to participate in the study. Nurses were not very optimistic about the usability of integrating audio recording in

clinical workflows as they perceived limited benefits in audio recording the HHC encounters. This is contrary to the findings of numerous studies [1-8] showing the importance of patient-spoken language in the development of diagnostic and risk identification algorithms for identifying patients with pathological entities and at risk of negative outcomes. Educating nurses about the importance of patients' spoken language may partially resolve this problem. In addition, nurses had concerns about increasing the length of HHC encounters because of the time required to set up the audio-recording device, which was a challenge given their existing heavy workload. The use of devices that are minimally burdensome and easy to use and enlisting the support of clinical managers are essential to address these issues and are key for successful integration.

Limitations

The findings of this study should be considered in light of several limitations. First, although the audio-recording devices selected for this study included a wide range of useful and convenient features for audio recording HHC patient-nurse encounters, they may not represent all features of existing devices in the market. For example, future studies may investigate the usability and functionality of Amazon Alexa (a device that was developed by the Amazon company) for audio recording patient-nurse encounters, the possibility of connecting to a secure server for storing the audio-recorded data, and the potential risks to patient privacy and confidentiality. Second, we quantified the quality of audio-recorded communication by measuring the accuracy of the automated transcription provided by AWS-GTS. Although the quality of audio-recorded data is correlated with the accuracy of automated transcription, it may not provide comprehensive insights into the quality of audio recorded by a device. Future studies may investigate other measures such as the sensitivity of the device for filtering background noise in noisy clinical settings to better evaluate the quality of the audio-recorded voice by the device. Third, as this study was conducted during the COVID-19 pandemic, recruitment and retention of HHC nurses were challenging. This was mostly because of nurses' heavy workload, and precautions needed to be taken to reduce the risk of COVID-19 transmission in the HHC setting. In addition, because of the COVID-19 pandemic, it was challenging for the research team to reach out to all nurses with a potential interest in participating in this study. Overall, we were able to recruit 5 nurses for the study of audio recording patient-nurse verbal communication conducted at VNSNY. Of the 5 nurses, 2 (40%) nurses left the VNSNY during the study. Hence, unfortunately, we could not reach the 2 nurses to solicit their perspectives. Although a sample size of 3 nurses may not provide sufficient data to achieve data saturation, the 3 nurses who participated in the interviews had extensive experience in HHC services, HHC workflow, and working with a racially diverse patient population in VNSNY. Therefore, they were able to provide a valuable evaluation of the facilitators and barriers to the pipeline designed for audio recording and its integration into the HHC workflow. Aggregation of the 3 nurses' perspectives sheds light on the facilitators and barriers to a large extent and was very informative for HHC managers and policy makers. For example, the manager of the VNSNY research center informed us that

they were willing to address some of the barriers (eg, training nurses about the importance of audio-recording encounters) to encourage more HHC nurses to participate in similar studies in the future. Currently, we are actively recruiting more nurses for this study, which will enable us to provide a better picture of nurses' opinions toward this process in our future report. Fourth, the racially diverse sample size of patients (10/45, 22%) who participated in this study was achieved through an iterative process and data saturation of thematic analysis of interview findings, which represents the patient population at the study agency to a large extent. However, the findings regarding patients' opinions may not provide a deep insight into the attitudes of ethnic minority patients or those with complex clinical conditions. Finally, this study was conducted at one agency, the largest HHC organization in the United States. However, there might be differences in the types of health care

services, communication, and practice patterns across HHC. Therefore, the overall findings of this study may not be representative of all HHC settings.

Conclusions

To develop an effective practical approach for integrating audio recording of patient-nurse verbal communication in HHC settings, it is essential to select an audio-recording device with high functionality and usability. Training nurses and clinical managers on the importance of audio-recorded verbal communication can encourage them to support the process of integration.

In addition, training can reduce the potential concerns of nurses about protecting their privacy and confidentiality during the recording process.

Conflicts of Interest

None declared.

Multimedia Appendix 1

Questions for semistructured interviews.

[[DOCX File, 98 KB - humanfactors_v9i2e35325_app1.docx](#)]

References

1. Kaddoura T, Vadlamudi K, Kumar S, Bobhate P, Guo L, Jain S, et al. Acoustic diagnosis of pulmonary hypertension: automated speech- recognition-inspired classification algorithm outperforms physicians. *Sci Rep* 2016 Sep 09;6:33182 [FREE Full text] [doi: [10.1038/srep33182](#)] [Medline: [27609672](#)]
2. Rao A, Huynh E, Royston TJ, Kornblith A, Roy S. Acoustic methods for pulmonary diagnosis. *IEEE Rev Biomed Eng* 2019;12:221-239. [doi: [10.1109/rbme.2018.2874353](#)]
3. Petti U, Baker S, Korhonen A. A systematic literature review of automatic Alzheimer's disease detection from speech and language. *J Am Med Inform Assoc* 2020 Nov 01;27(11):1784-1797 [FREE Full text] [doi: [10.1093/jamia/ocaa174](#)] [Medline: [32929494](#)]
4. Low D, Bentley K, Ghosh S. Automated assessment of psychiatric disorders using speech: a systematic review. *Laryngoscope Investig Otolaryngol* 2020 Feb;5(1):96-116 [FREE Full text] [doi: [10.1002/lio2.354](#)] [Medline: [32128436](#)]
5. Yadav S, Keerthana M, Gope D, Maheswari K U, Ghosh PK. Analysis of acoustic features for speech sound based classification of asthmatic and healthy subjects. In: Proceedings of the ICASSP 2020 - 2020 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP). 2020 Presented at: ICASSP 2020 - 2020 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP); May 4-8, 2020; Barcelona, Spain. [doi: [10.1109/icassp40776.2020.9054062](#)]
6. Cohen AS, Cox CR, Le TP, Cowan T, Masucci MD, Strauss GP, et al. Using machine learning of computerized vocal expression to measure blunted vocal affect and alogia. *NPJ Schizophr* 2020 Sep 25;6(1):26 [FREE Full text] [doi: [10.1038/s41537-020-00115-2](#)] [Medline: [32978400](#)]
7. Pramanik M, Pradhan R, Nandy P, Qaisar SM, Bhoi AK. Assessment of acoustic features and machine learning for Parkinson's detection. *J Healthc Eng* 2021 Aug 21;2021:9957132-9957113 [FREE Full text] [doi: [10.1155/2021/9957132](#)] [Medline: [34471507](#)]
8. Mahajan P, Baths V. Acoustic and language based deep learning approaches for Alzheimer's dementia detection from spontaneous speech. *Front Aging Neurosci* 2021 Feb 5;13:623607 [FREE Full text] [doi: [10.3389/fnagi.2021.623607](#)] [Medline: [33613269](#)]
9. Barr P, Bonasia K, Verma K, Dannenberg M, Yi C, Andrews E, et al. Audio-/videorecording clinic visits for patient's personal use in the United States: cross-sectional survey. *J Med Internet Res* 2018 Sep 12;20(9):e11308 [FREE Full text] [doi: [10.2196/11308](#)] [Medline: [30209029](#)]
10. Meeusen A, Porter R. Patient-reported use of personalized video recordings to improve neurosurgical patient-provider communication. *Cureus* 2015 Jun;7(6):e273 [FREE Full text] [doi: [10.7759/cureus.273](#)] [Medline: [26180697](#)]
11. Ball SL, Weiner SJ, Schwartz A, Altman L, Binns-Calvey A, Chan C, et al. BMC Health Serv Res 2021 Aug 30;21(1):891 [FREE Full text] [doi: [10.1186/s12913-021-06921-3](#)] [Medline: [34461903](#)]

12. Romagnoli K, Handler S, Hochheiser H. Home care: more than just a visiting nurse. *BMJ Qual Saf* 2013 Dec;22(12):972-974 [[FREE Full text](#)] [doi: [10.1136/bmjqs-2013-002339](https://doi.org/10.1136/bmjqs-2013-002339)] [Medline: [23940375](https://pubmed.ncbi.nlm.nih.gov/23940375/)]
13. Shang J, Russell D, Dowding D, McDonald MV, Murtaugh C, Liu J, et al. A predictive risk model for infection-related hospitalization among home healthcare patients. *J Healthc Qual* 2020;42(3):136-147. [doi: [10.1097/jhq.0000000000000214](https://doi.org/10.1097/jhq.0000000000000214)]
14. Jones C, Bowles K, Richard A, Boxer R, Masoudi F. High-value home health care for patients with heart failure. *Circ Cardiovascular Quality Outcomes* 2017 May;10(5):e003676. [doi: [10.1161/circoutcomes.117.003676](https://doi.org/10.1161/circoutcomes.117.003676)]
15. Kortum P, Peres S. Evaluation of home health care devices: remote usability assessment. *JMIR Hum Factors* 2015 Jun 05;2(1):e10 [[FREE Full text](#)] [doi: [10.2196/humanfactors.4570](https://doi.org/10.2196/humanfactors.4570)] [Medline: [27025664](https://pubmed.ncbi.nlm.nih.gov/27025664/)]
16. Weinger M, Wiklund M, Gardner-Bonneau D. *Handbook of Human Factors in Medical Device Design*. Boca Raton, Florida: CRC Press; 2010.
17. Barr PJ, Dannenberg MD, Ganoe CH, Haslett W, Fail R, Hassanpour S, et al. Sharing annotated audio recordings of clinic visits with patients-development of the open recording automated logging system (ORALS): study protocol. *JMIR Res Protoc* 2017 Jul 06;6(7):e121 [[FREE Full text](#)] [doi: [10.2196/resprot.7735](https://doi.org/10.2196/resprot.7735)] [Medline: [28684387](https://pubmed.ncbi.nlm.nih.gov/28684387/)]
18. Barr P, Haslett W, Dannenberg M, Oh L, Elwyn G, Hassanpour S, et al. An audio personal health library of clinic visit recordings for patients and their caregivers (HealthPAL): user-centered design approach. *J Med Internet Res* 2021 Oct 22;23(10):e25512 [[FREE Full text](#)] [doi: [10.2196/25512](https://doi.org/10.2196/25512)] [Medline: [34677131](https://pubmed.ncbi.nlm.nih.gov/34677131/)]
19. Sedgwick P, Greenwood N. Understanding the Hawthorne effect. *BMJ* 2015 Sep 04;351:h4672. [doi: [10.1136/bmj.h4672](https://doi.org/10.1136/bmj.h4672)] [Medline: [26341898](https://pubmed.ncbi.nlm.nih.gov/26341898/)]
20. Borsci S, Federici S, Lauriola M. On the dimensionality of the System Usability Scale: a test of alternative measurement models. *Cogn Process* 2009 Aug;10(3):193-197. [doi: [10.1007/s10339-009-0268-9](https://doi.org/10.1007/s10339-009-0268-9)] [Medline: [19565283](https://pubmed.ncbi.nlm.nih.gov/19565283/)]
21. Vogel A, Morgan A. Factors affecting the quality of sound recording for speech and voice analysis. *Int J Speech Language Path* 2009 Dec;11(6):431-437. [doi: [10.1080/17549500902822189](https://doi.org/10.1080/17549500902822189)]
22. Schneider S, Baevski A, Collobert R, Auli M. wav2vec: Unsupervised pre-training for speech recognition. *Proc Interspeech* 2019. URL: https://www.isca-speech.org/archive/interspeech_2019/schneider19_interspeech.html [accessed 2022-04-10]
23. Riegel B, Dickson VV, Topaz M. Qualitative analysis of naturalistic decision making in adults with chronic heart failure. *Nurs Res* 2013;62(2):91-98. [doi: [10.1097/NNR.0b013e318276250c](https://doi.org/10.1097/NNR.0b013e318276250c)] [Medline: [23154659](https://pubmed.ncbi.nlm.nih.gov/23154659/)]
24. Topaz M, Bar-Bachar O, Admi H, Denekamp Y, Zimlichman E. Patient-centered care via health information technology: a qualitative study with experts from Israel and the U.S. *Inform Health Soc Care* 2020 Sep 27;45(3):217-228. [doi: [10.1080/17538157.2019.1582055](https://doi.org/10.1080/17538157.2019.1582055)] [Medline: [30917717](https://pubmed.ncbi.nlm.nih.gov/30917717/)]
25. Topaz M, Ronquillo C, Peltonen LM, Pruinelli L, Sarmiento RF, Badger MK, et al. Nurse informaticians report low satisfaction and multi-level concerns with electronic health records: results from an international survey. *AMIA Annu Symp Proc* 2016;2016:2016-2025 [[FREE Full text](#)] [Medline: [28269961](https://pubmed.ncbi.nlm.nih.gov/28269961/)]

Abbreviations

AWS-GTS: Amazon Web Service General Transcribe System

EHR: electronic health record

HHC: home health care

RA: research assistant

SUS: System Usability Scale

VNSNY: Visiting Nurse Service of New York

WER: word error rate

Edited by A Kushniruk; submitted 01.12.21; peer-reviewed by C Eldredge, R Pryss; comments to author 20.01.22; revised version received 20.02.22; accepted 21.03.22; published 11.05.22.

Please cite as:

Zolnoori M, Vergez S, Kostic Z, Jonnalagadda SR, V McDonald M, Bowles KKH, Topaz M

Audio Recording Patient-Nurse Verbal Communications in Home Health Care Settings: Pilot Feasibility and Usability Study

JMIR Hum Factors 2022;9(2):e35325

URL: <https://humanfactors.jmir.org/2022/2/e35325>

doi: [10.2196/35325](https://doi.org/10.2196/35325)

PMID: [35544296](https://pubmed.ncbi.nlm.nih.gov/35544296/)

©Maryam Zolnoori, Sasha Vergez, Zoran Kostic, Siddhartha Reddy Jonnalagadda, Margaret V McDonald, Kathryn K H Bowles, Maxim Topaz. Originally published in *JMIR Human Factors* (<https://humanfactors.jmir.org>), 11.05.2022. This is an open-access article distributed under the terms of the Creative Commons Attribution License (<https://creativecommons.org/licenses/by/4.0/>),

which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Human Factors, is properly cited. The complete bibliographic information, a link to the original publication on <https://humanfactors.jmir.org>, as well as this copyright and license information must be included.

Letter to the Editor

Should Expert Surgeon Guidance Be Given Remotely? Comment on “User Experience in Remote Surgical Consultation: Survey Study of User Acceptance and Satisfaction in Real-Time Use of a Telemedicine Service”

Daniella Soussi^{1*}, BSc; Chiara Jade Vedi^{1*}, BSc

School of Medicine, Imperial College London, London, United Kingdom

* all authors contributed equally

Corresponding Author:

Daniella Soussi, BSc

School of Medicine

Imperial College London

Exhibition Rd

South Kensington

London, SW7 2BX

United Kingdom

Phone: 44 02075895111

Email: d.soussis@yahoo.gr

Related Article:

Comment on: <https://humanfactors.jmir.org/2021/4/e30867>

(*JMIR Hum Factors* 2022;9(2):e36681) doi:[10.2196/36681](https://doi.org/10.2196/36681)

KEYWORDS

telemedicine; user experience; satisfaction; technology acceptance; usability; perioperative; surgery; consultation; surgeons; performance; evaluation; tele-guidance; telehealth; telemedicine implementation; telementoring; surgical consultation; usefulness

We recently read with great interest the article “User Experience in Remote Surgical Consultation: Survey Study of User Acceptance and Satisfaction in Real-Time Use of a Telemedicine Service” by Aminoff et al [1], where a real-time telemedicine service was used for intraoperative surgical consultation assistance during endoscopic retrograde cholangiopancreatography procedures. The authors investigated the surgeons’ preprocedure expectations of how the service would aid operative performance and patient outcomes, and their satisfaction after use of the service. We commend the authors for their work; however, we would like to highlight some caveats in their research.

Although this study presents interim results, we are concerned that the perceived future use of the service may be overestimated. Expert surgeons had to book out their time in advance to participate, and when the sample was asked about the perceived demand for teleguidance, 80% said they believed there would be no demand for it or were unsure of the demand [1]. This coupled with the fact that sites were encouraged to use the service due to being part of a trial makes it unclear as to whether the telemedicine service would be used in practice.

Moreover, technical issues experienced in virtual services can prohibit their successful implementation [2]. Technical issues were apparent in 24% of cases, and problems in audio and video connection and fluoroscopy transfer could lead to incorrect surgical decision-making guidance from the remote expert. Even though some of these issues were resolved either through connection restarting or calling medical technicians, these unnecessary delays increase patients’ time under anesthesia, which is known to increase perioperative complications such as infection [3]. Indeed, this could explain the 4 (2.8%) cases that had postoperative complications; although, we cannot confirm this as a causal inference [1]. Furthermore, due to the limitations in the content being transferred at any one time via the service and the lack of 3D visualization, we caution against the use of remote expert guidance in particularly difficult surgical decision-making scenarios.

Although novice endoscopic retrograde cholangiopancreatographers demonstrated an increase in their level of expertise during the study period, we argue that this intervention may be time-inefficient, with 81.1% of the procedure time being spent on receiving guidance [1].

Furthermore, no follow-up data is available to predict whether skills would be maintained or regress if tele-guidance were not implemented, which casts doubt on the service's training value.

Previous work has found that *telestration* features can reduce the time of mentoring sessions [4]. However, due to the user-unfriendly design of this particular *telestration* tool, this benefit was unlikely. Therefore, we suggest the development of a more tailored design and function to create more time for other clinical duties consulting surgeons may have.

Finally, to determine the true value of this service in clinical practice, we recommend a study investigating the acquisition of both technical and nontechnical skills for on-site expert-guided surgeries compared to remote surgical consultation.

In conclusion, we commend the authors for their innovative surgical guidance service and suggest that the aforementioned are considered when formally implementing the service in the clinical setting.

Conflicts of Interest

None declared.

Editorial Notice

The corresponding author of "User Experience in Remote Surgical Consultation: Survey Study of User Acceptance and Satisfaction in Real-Time Use of a Telemedicine Service" declined to respond to this letter.

References

1. Aminoff H, Meijer S, Groth K, Arnelo U. User experience in remote surgical consultation: survey study of user acceptance and satisfaction in real-time use of a telemedicine service. *JMIR Hum Factors* 2021 Nov 30;8(4):e30867 [FREE Full text] [doi: [10.2196/30867](https://doi.org/10.2196/30867)] [Medline: [34851302](https://pubmed.ncbi.nlm.nih.gov/34851302/)]
2. Greenhalgh T, Vijayaraghavan S, Wherton J, Shaw S, Byrne E, Campbell-Richards D, et al. Virtual online consultations: advantages and limitations (VOCAL) study. *BMJ Open* 2016 Jan 29;6(1):e009388 [FREE Full text] [doi: [10.1136/bmjopen-2015-009388](https://doi.org/10.1136/bmjopen-2015-009388)] [Medline: [26826147](https://pubmed.ncbi.nlm.nih.gov/26826147/)]
3. Ravi B, Jenkinson R, O'Heireamhoin S, Austin PC, Aktar S, Leroux TS, et al. Surgical duration is associated with an increased risk of periprosthetic infection following total knee arthroplasty: a population-based retrospective cohort study. *EClinicalMedicine* 2019 Nov;16:74-80 [FREE Full text] [doi: [10.1016/j.eclinm.2019.09.015](https://doi.org/10.1016/j.eclinm.2019.09.015)] [Medline: [31832622](https://pubmed.ncbi.nlm.nih.gov/31832622/)]
4. Budrionis A, Hasvold P, Hartvigsen G, Bellika JG. Assessing the impact of telestration on surgical telementoring: a randomized controlled trial. *J Telemed Telecare* 2016 Jan;22(1):12-17. [doi: [10.1177/1357633X15585071](https://doi.org/10.1177/1357633X15585071)] [Medline: [26026177](https://pubmed.ncbi.nlm.nih.gov/26026177/)]

Edited by T Leung; submitted 22.01.22; peer-reviewed by D Gunasekeran; accepted 09.06.22; published 29.06.22.

Please cite as:

Soussi D, Vedi CJ

Should Expert Surgeon Guidance Be Given Remotely? Comment on "User Experience in Remote Surgical Consultation: Survey Study of User Acceptance and Satisfaction in Real-Time Use of a Telemedicine Service"

JMIR Hum Factors 2022;9(2):e36681

URL: <https://humanfactors.jmir.org/2022/2/e36681>

doi: [10.2196/36681](https://doi.org/10.2196/36681)

PMID: [35767326](https://pubmed.ncbi.nlm.nih.gov/35767326/)

©Daniella Soussi, Chiara Jade Vedi. Originally published in *JMIR Human Factors* (<https://humanfactors.jmir.org>), 29.06.2022. This is an open-access article distributed under the terms of the Creative Commons Attribution License (<https://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in *JMIR Human Factors*, is properly cited. The complete bibliographic information, a link to the original publication on <https://humanfactors.jmir.org>, as well as this copyright and license information must be included.

Original Paper

Application of a Web-based Self-assessment Triage Tool During the COVID-19 Pandemic: Descriptive Study

Anna Nowicka^{1,2*}, MD; Jakub Jaszczak^{2*}, MD; Anna Szymanek Pasternak^{1,3}, MD, PhD; Krzysztof Simon^{1,3}, MD, PhD

¹Provincial Specialist Hospital them. J. Gromkowski, Wrocław, Poland

²Infermedica, Wrocław, Poland

³Department of Infectious Diseases and Hepatology, Wrocław Medical University, Wrocław, Poland

*these authors contributed equally

Corresponding Author:

Anna Nowicka, MD

Provincial Specialist Hospital them. J. Gromkowski

Koszarowa 5

Wrocław, 51-149

Poland

Phone: 48 660427576

Email: nowickanna@gmail.com

Abstract

Background: The COVID-19 pandemic has sped up the implementation of telehealth solutions in medicine. A few symptom checkers dedicated for COVID-19 have been described, but it remains unclear whether and how they can affect patients and health systems.

Objective: This paper demonstrates our experiences with the COVID-19 risk assessment (CRA) tool. We tried to determine who the user of the web-based COVID-19 triage app is and compare this group with patients in the infectious diseases ward's admission room to evaluate who could benefit from implementing the COVID-19 online symptom checker as a remote triage solution.

Methods: We analyzed the answers of 248,862 people interacting with an online World Health Organization–based triage tool for assessing the probability of SARS-CoV-2 infection. These users filled in an online questionnaire between April 7 and August 6, 2020. Based on the presented symptoms, risk factors, and demographics, the tool assessed whether the user's answers were suggestive of COVID-19 and recommended appropriate action. Subsequently, we compared the sociodemographic and clinical characteristics of tool users with patients admitted to the Infectious Diseases Admission Room of J. Gromkowski Hospital in Wrocław.

Results: The CRA tool tended to be used by asymptomatic or oligosymptomatic individuals (171,226 [68.80%] of all users). Most users were young (162,432 [65.27%] were below 40 years of age) and without comorbidities. Only 77,645 (31.20%) of the self-assessment app users were suspected of COVID-19 based on their reported symptoms. On the contrary, most admission room patients were symptomatic—symptoms such as fever, cough, and dyspnea were prevalent in both COVID-19-positive and COVID-19-negative patients. COVID-19-suspected patients in the CRA tool group presented similar COVID-19 symptoms as those who presented to the admission room. These were cough (25,062/40,007 [62.64%] in the CRA tool group vs 138/232 [59.48%] in the admission room group), fever (23,123/40,007 [57.80%] in the CRA tool group vs 146/232 [62.93%] in the admission room group), and shortness of breath (15,157/40,007 [37.89%] in the CRA tool group vs 87/232 [37.50%] in the admission room group).

Conclusions: The comparison between the symptomatology of the users interacting with the CRA tool and those visiting the admission room revealed 2 major patient groups who could have benefited from the implementation of the self-assessment app in preclinical triage settings. The primary users of the CRA tool were young, oligosymptomatic individuals looking for screening for COVID-19 and reassurance early in the COVID-19 pandemic. The other group were users presenting the typical symptoms suggestive of COVID-19 at that time. The CRA tool recognized these individuals as potentially COVID-19 positive and directed them to the proper level of care. These use cases fulfil the idea of preclinical triage; however, the accuracy and influence on health care must be examined in the clinical setting.

KEYWORDS

COVID-19; symptom checker; preclinical triage; self-assessment tool; online applications; COVID-19 remote triage; COVID-19 self-assessment

Introduction

Background

After the outbreak of the COVID-19 pandemic, the health care systems of affected countries faced an unprecedented challenge. Ensuring the continuity of care and screening the vast number of suspected patients have put a significant strain on health care, leading to the depletion of public health resources [1,2]. Although the health system resources were transferred to provide critical services to patients suffering from COVID-19, the utilization of medical visits reduced by even 42% [2], suggesting that patients with less severe illnesses tended to avoid in-person consultation or had no possibility to attend one.

During the pandemic, especially in the early days, there was much uncertainty regarding the symptomatology and clinical course of the novel coronavirus disease. This has been reflected in the number of searches for the phrase “covid 19 symptoms” on the Google platform, which at the time of the study varied from 443,000 to 2.2 million searches per month just for the United States [3].

These uncertain times have presented an opportunity to popularize telehealth solutions in medicine. The means of remote consultations have found their way mostly in primary care as a substitute for in-person visits [4] but also as a way of remote triage of COVID-19 patients.

Triage is defined as a classification of patients according to their urgencies. Remote triage uses the means of distance communication, such as telephones or interactive websites, allowing for the segregation of patients before they interact with health care professionals. Remote triage solutions have been proven helpful in telephone call centers, where they have been associated with lower in-person health care use [5]. They have also been demonstrated to be useful in the triage of COVID-19 patients, as they have reduced the number of unnecessary consultations, hence reducing the exposure of the staff to COVID-19 [6]. Web-based COVID-19 symptom checkers and triage tools have also proved useful in scheduling tests [7,8], monitoring symptoms [9-11], providing evidence-based educational value [8,9,12], and supporting self-isolation [13].

Objective

In this study, we wanted to share our findings regarding the COVID-19 risk assessment (CRA) tool. It was a World Health Organization (WHO) guidelines-based online triage tool, which assessed the risk of SARS-CoV-2 infection and returned a probable outcome with a concise recommendation of what to do next, along with evidence-based educational materials about COVID-19.

We gathered and analyzed the data of 651,757 patients interacting with the CRA tool, focusing on their demographics, risk factors, reported symptoms, possible exposure to

SARS-CoV-2, and recommended triage. The aim was to establish who the main users of web-based COVID-19 symptom checkers (age, sex, comorbidities, presenting symptoms) are and who might have benefitted from implementing COVID-19 symptom checkers as preclinical triage solutions.

Since confirming the diagnosis in an online self-assessment tool was not achievable, we compared the results (sociodemographic and clinical characteristics of CRA users) with the health records of the Infectious Diseases Admission Room of J. Gromkowski Hospital in Wrocław to establish whether and how these groups corresponded. The goal was to evaluate who could benefit from implementing this solution as preclinical triage.

Methods

Study Population

Since April 7, 2020, we have been collecting and utilizing responses from the CRA tool users. The app was developed by Infermedica company, as a non-profit project. It utilized a diagnostic algorithm designed based on WHO and Centers for Disease Control and Prevention (CDC) recommendations. The specific time frame was chosen due to periodical updates of the app questions flow. In the selected period, there were no major changes to the question flow so that the collected information could be unbiased.

Inclusion Criteria

The study population included individuals concerned about their risk of COVID-19 infection:

- Users who filled the questionnaire available through the Infermedica website between April 7 and August 6, 2020
- Users who filled the questionnaires available on third-party websites, which obtained permission to use our tool within their platforms between April 7 and August 6, 2020

Exclusion Criteria

The exclusion criteria were:

- Completing the interview in an outdated 1.0 and 2.0 version (not all providers of our tool updated their software before the beginning of the study)
- Completing the interview in a version customized for a national health system so that it was incompatible with WHO and CDC recommendations
- Not completing the whole interview
- Age below 18 years
- Completing the interview in a language other than Polish

Data Privacy and Ethical Statement

The study population consisted of 2 arms: users of the web app and patients in the admission room.

The app arm consisted of users of the web app who accepted the terms of service. All data processed through the COVID-19 risk assessment checker were anonymous and did not allow us to identify an individual based on the information provided during the interview. Informed consent to use anonymized data was provided by the users by accepting the terms of service. A privacy policy and personal data protection were applied.

The admission room arm of the study did not require ethics committee approval as a retrospective study, according to the guidelines of the local ethical compliance body [14].

COVID-19 Symptom Checker Characteristics

The CRA is a triage tool dedicated to nonprofessional users. The checkup was designed to assess whether the user's symptoms may be the result of SARS-CoV-2 infection. It had a form of a responsive web app that could be embedded within a website or an Application Program Interface (API) that can serve as a technological core for building custom apps. (An API is a set of routines, protocols, and tools for building software applications. Basically, an API specifies how software components should interact. It serves as a technological core for custom-building applications.) The flow of the interview was solely based on the official WHO guidelines for diagnosing COVID-19 [15]. The first version of the API was released on March 20, 2020 (version 1.0), followed by updates on March 25, 2020 (version 2.0), April 7, 2020 (version 3.0), and May 7, 2020 (version 4.0).

The app has been considered final from version 3.0; the set of risk factors and symptoms have reached their final form. However, the core logic of the interview, such as the flow of the interview, types of acquired data, and types of output recommendations, has been consistent from the first released version. In this study, we only considered interviews in the period between April 7 and August 6, 2020.

Medical Foundation

The CRA tool's logic was built around WHO guidelines [15] and WHO daily transmission reports [16]. The interview was designed to gather enough data to establish whether the user falls into any of the 3 categories mentioned in said guidelines as "Suspected case" for COVID-19; therefore, the reported symptoms may have resulted from SARS-CoV-2 infection.

For this reason, the interview consisted of 3 sets of questions that could be grouped into 3 categories:

- Risk factors and symptoms
- Places of residence and travel
- Contact with possible COVID-19 cases

In some cases, when this information was unnecessary to make a diagnosis, some questions were omitted.

Data Analysis

The majority of the data were compared and presented with the use of descriptive statistics. Inferential statistics had to be omitted because of the significant differences in both compared populations and vastly different sample sizes. We decided to only use statistical analysis to compare comorbidities related to COVID-19 in both CRA and admission room groups. In

CRA, *P* values were calculated with the test of proportions and in the admission room, with the Fisher exact test.

Screen Deep Dive

The interview consisted of up to 8 consecutive screens. Not every screen had to be included; this is the maximum number of screens that the user could have been exposed to. If the patient reported emergency evidence (ie, acute dyspnea), the interview was terminated with an instruction to call an ambulance. The screens in the display order were "Welcome & Terms of Service," "Age and Sex Selection," "Risk Factors," "Symptoms," "Red Flags," "Possible Exposure to COVID-19," "Travel and Residency," and "Outcome."

Nine risk factors were included to inquire about the user's chronic illnesses and overall medical condition: diseases or drugs that weaken the immune system, obesity, long-term stay at a care facility or nursing home, diabetes, cancer, cardiovascular disease, history of chronic lung disease, history of chronic liver disease, and history of chronic kidney disease.

Some of these comorbidities have been described as negatively impacting COVID-19 infection outcomes [17]. We also included risk factors described in the Pneumonia Severity Index (PSI) as a negative prognostic factor indicating the need for hospitalization [18].

The symptom screens were oriented on inquiring about users' symptoms that should raise clinical suspicion for COVID-19 according to WHO guidelines [15]. There was a list of 11 symptoms users could choose from: fever, cough, shortness of breath, fatigue, muscle pain, chills, headache, diarrhea, nausea, sore throat, and impaired taste or smell.

Furthermore, the interview focused on assessing red flags—immediate health threats to the user that should yield in cessation of the interview. To do so, the user was asked about rapid symptom deterioration, tachypnea, or hemoptysis.

There were 6 possible outcomes of the interview, which referred to the possibility of COVID-19 infection and the severity of symptoms:

- COVID-19 suspected, serious: "Call the emergency number. Avoid all contact."
- COVID-19 suspected, nonserious: "Consult your health care provider. Avoid all contact."
- Contact with COVID-19, no symptoms: "Quarantine."
- Non-COVID-19, serious: "Call a doctor."
- Non-COVID-19, nonserious: "Stay home and monitor your symptoms."
- Asymptomatic: "Follow preventive measures."

The extensive screen description and decision tree logic can be browsed in [Multimedia Appendix 1](#).

Comparison Group: Admission Room Analysis

To compare individuals completing the survey with real patients diagnosed with COVID-19 by health care professionals, we turned to the Infectious Diseases Admission Room of J. Gromkowski Hospital in Wrocław. We analyzed 291 cases of patients visiting the admission room between April 7 and August 6, 2020. All the patients reporting to the admission room were

suspected of COVID-19 infection; no other cases of infectious diseases were consulted in the admission room at that time. They may have been brought to the admission room by ambulance, referred by the primary care physician, or admitted by themselves. We excluded patients below 18 years of age.

Each patient was interviewed and examined by the physician working in the admission room. The interview consisted of fixed elements, such as current symptoms, comorbidities, medication, history of travel, contact with COVID-19-positive persons, and workplace and family interview. Blood analysis, chest X-rays, and COVID-19 swabs were obtained in most cases.

The patient's history and examination, along with the additional tests, allowed them to decide on admission to the hospital or discharge. After 24 hours, the results of the COVID-19 genetic test (reverse transcription polymerase chain reaction [RT-PCR] from nasopharyngeal or pharyngeal swabs) were available, which allowed reaching the final diagnosis.

Setting

J. Gromkowski Hospital in Wrocław, Lower Silesian Voivodeship, Poland, is 1 of the specialist hospitals in that city. There are 2 infectious disease wards in the hospital. The Infectious Diseases Admission Room serves as the place for preliminary triage, diagnosis, and treatment of incoming patients suspected of contracting infectious diseases. During the COVID-19 pandemic, it served as the main consultation facility of COVID-19-suspected cases.

Population

In this study, we analyzed the Infectious Diseases Admission Room cases between April 7 and August 6, 2020. We focused on the set of reported symptoms, comorbidities, contact with COVID-19 cases, and travel history. Our goal was to determine the patient profile, meaning assessing the set of symptoms connected with COVID-19 cases compared to non-COVID-19 cases.

Finally, we wanted to compare the sociodemographic and clinical characteristics of hospital patients and the ones completing the self-assessment interview.

Symptoms

In the study, we screened for 8 symptoms that are suggestive of COVID-19 infection: cough, fever, dyspnea, diarrhea, myalgia, rhinorrhea, taste and smell abnormalities, and pharyngeal pain.

Results

Demographics and Groups Characteristics

CRA Tool

Of the 697,903 individual interviews performed on the CRA tool between April 7 and August 6, 2020, a total of 248,862 (35.66%) individual interviews met the inclusion criteria. Most of these interviews came from the government portal of the Polish Ministry of Health, which embedded the app within its website [19]: 117,311 (47.14%) of all interviews. In addition, 91,805 (36.89%) interviews were performed on the original CRA website [20], and 17,767 (7.14%) interviews were performed on the COVID-19 mobile app commissioned by the Polish Ministry of Health. Other notable institutions adopting the CRA tool and providing us interviews analyzed in the study included PZU Zdrowie (Polish biggest private health care provider), Dovera (private health care provider in Slovakia), Global Excel (medical assistance company operating in the U.S. and Canada), and others [21]. The CRA tool is offered in 37 languages in total: Polish, English, Slovak, Ukrainian, Portuguese-Brazilian, and Russian are the most popular languages. However, only Polish-speaking users met the inclusion criteria (Figure 1).

Most of the respondents were between 18 and 40 years old ($n=158,998$ [63.89%] of all respondents). The least prevalent were users between 80 and 90 years old ($n=498$, 0.2%). The mean age was 37 years. The study included 130,966 (52.63%) males and 117,896 (47.37%) females (Figure 2).

Figure 1. Finished CRA interviews daily (blue line); for comparison, daily number of new diagnosed COVID-19 cases in Poland (red line). CRA: COVID-19 risk assessment.

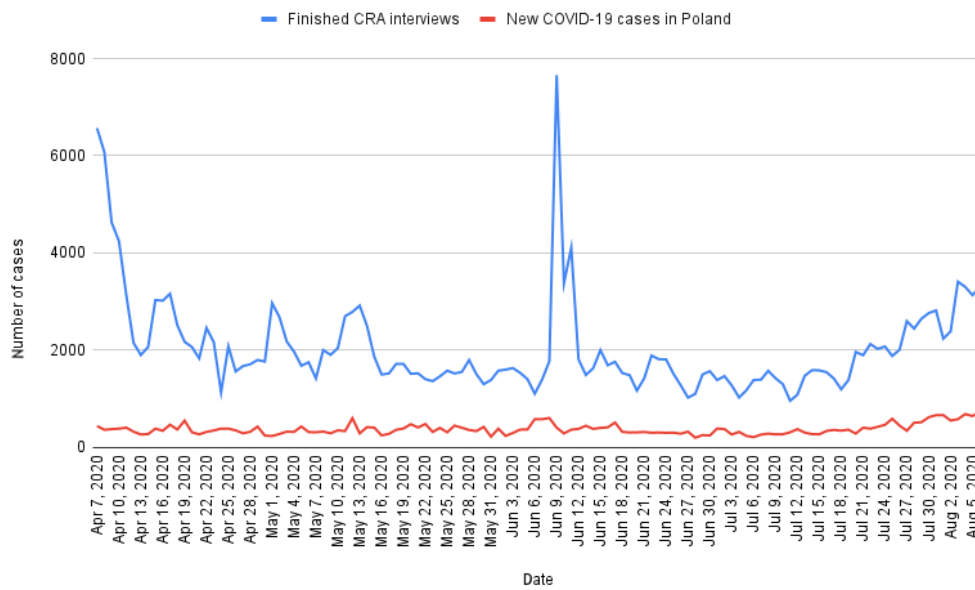
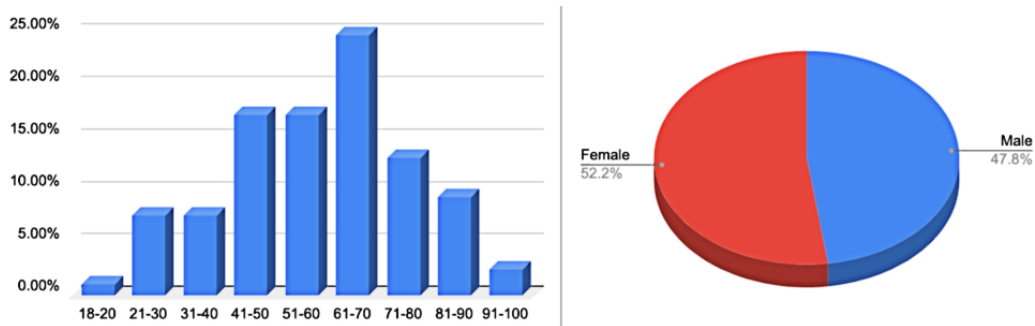


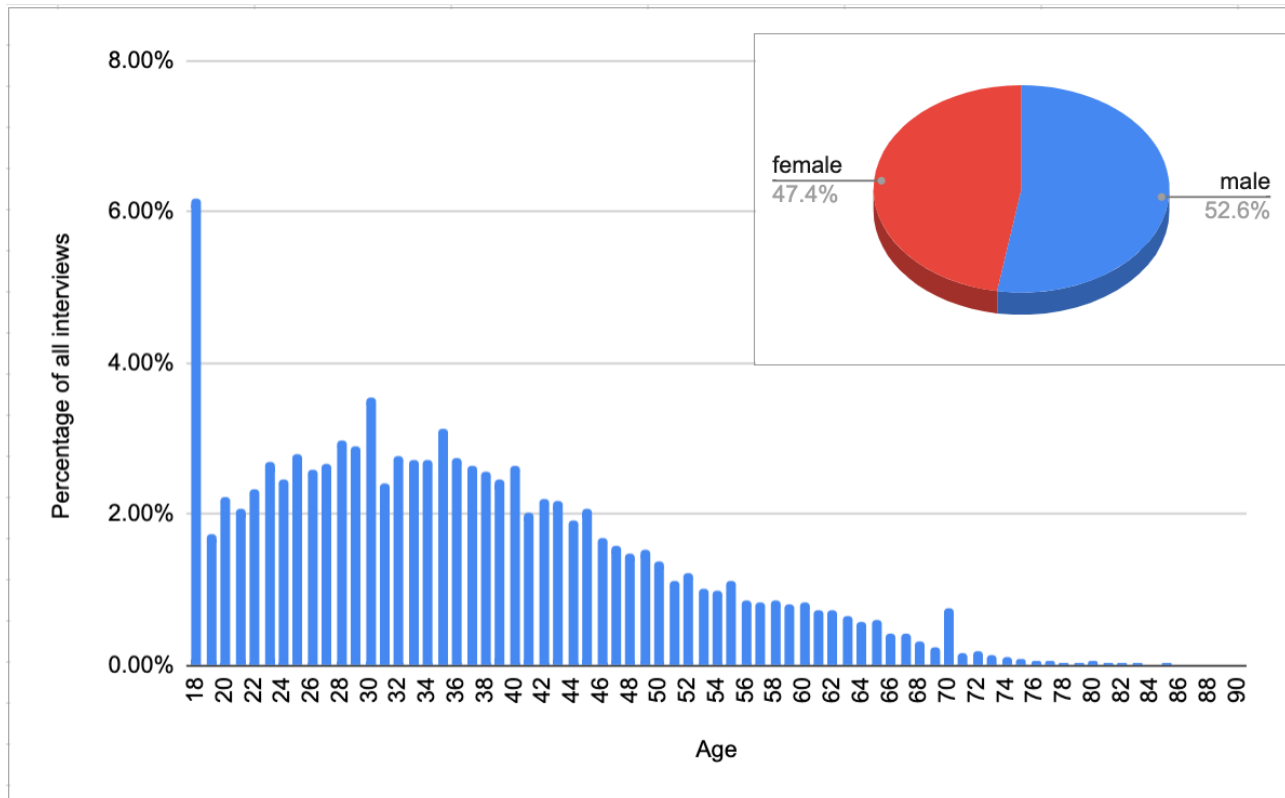
Figure 2. Age and sex distribution of admission room patients (N=291).



Admission Room

The study included 291 patients who visited the Infectious Diseases Admission Room of J. Gromkowski Hospital in Wrocław between April 7 and August 6, 2020. There were 152

(52.23%) women and 139 (47.77%) men enrolled in the study. Most of the patients were between 41 and 70 years old. The mean age was 58 years; the median age was 60 years (Figure 3).

Figure 3. Age and sex distribution of CRA users. CRA: COVID-19 risk assessment.

Outcomes and Triage Results

CRA Tool

Among the users of the CRA tool, the most common interview result was “asymptomatic” or “Follow preventive measures,” which was displayed to 98,081 (39.41%) of the 248,862 users. This subgroup consisted of users who answered the questionnaire but denied having any symptoms or COVID-19 exposure.

The second-most common triage outcome was “non-COVID-19, nonserious” or “Stay home and monitor your symptoms” for 73,145 (29.39%) of the 248,862 users. This subgroup comprised users who answered the questionnaire and reported only mild symptoms, such as fatigue, muscle pain, chills, headache, diarrhea, nausea, sore throat, and impaired taste or smell, but denied having any COVID-19 exposure (contact or travel). These users were not suspected of SARS-CoV-2 infection according to the diagnosing rules proposed by WHO at that time [15].

Both these groups added up to 171,226 (68.80%), which made them the majority of the CRA tool users. See [Tables 1](#) and [2](#) for details of the CRA tool group.

The third-most common triage outcome was “Call the emergency number,” which was recommended to 30,494 (12.25%) of the users. These were referred to as “COVID-19 suspected, serious” cases. Users who received that recommendation reported red-flag symptoms indicating respiratory distress or potentially severe infection (shortness of breath in the elderly, tachypnea, hemoptysis, high-grade fever, rapid symptom deterioration) and confirmed potential COVID-19 exposure.

Of the 248,862 users, 21,980 (8.83%) were classified as “Non-COVID-19, serious”: they received a “Call a doctor” recommendation. These users were not suspected of SARS-CoV-2 infection, because they had not met the WHO criteria of the suspected case at the time [15], but were advised to obtain a teleconsultation due to potentially severe symptoms: shortness of breath, high-grade fever, and fever and cough in the elderly.

The least prevalent group was the “COVID-19 suspected, nonserious” or “Consult your health care provider. Avoid all contact,” displayed to 9513 (3.82%) users. This group reported symptoms and COVID-19 exposure suggestive of SARS-CoV-2 infection but denied having potentially serious symptoms calling for an in-person consultation. They were advised to self-isolate and undergo a COVID-19 test.

Table 1. Distribution of CRA^a interview outcomes (N=248,862).

Triage	Patients, n (%)
Asymptomatic	98,081 (39.41)
Non-COVID-19, nonserious	73,145 (29.39)
COVID-19 suspected, serious	30,494 (12.25)
Non-COVID-19, serious	21,980 (8.83)
Quarantine	15,649 (6.29)
COVID-19 suspected, nonserious	9513 (3.82)

^aCRA: COVID-19 risk assessment.

Table 2. Distribution of CRA^a interview outcomes: matrix of the clinical suspicion of COVID-19 (N=248,862).

Severity of presented symptoms	COVID-19 suspected, n (%)	Non-COVID-19, n (%)
Serious	30,494 (12.25)	21,980 (8.83)
Nonserious	9513 (3.82)	171,226 (68.80)

^aCRA: COVID-19 risk assessment.

Admission Room

Of the 291 patients, 232 (79.73%) tested positive for COVID-19 and 59 (20.27%) tested negative for COVID-19. Of the 152 women, 126 (82.89%) were COVID-19 positive and 26 (17.11%) were COVID-19 negative. Of the 139 men, 106 (76.26%) were COVID-19 positive and 33 (23.74%) were COVID-19 negative.

Most of the patients (n=167, 57.39%) of the admission room group were classified by consulting physicians as patients in good general condition, 85 (29.21%) of the patients were judged to be in moderate general condition, 30 (10.31%) were in a bad general condition, and 9 (3.09%) were in a severely bad general condition.

Comorbidities

The number of reported comorbidities in the self-assessment app was 71,515; at least 1 risk factor was reported in 71,523 (28.74%) of the interviews. In other words, in 177,339 (71.26%) of the interviews, users did not report any comorbidity.

The most frequently reported comorbidity in both the CRA tool users and the admission room patients was cardiovascular disease, defined as hypertension, coronary disease, or heart insufficiency and confirmed by 37,628 (15.12%) of 248,862

CRA tool users and 138 (47.42%) of 291 admission room patients.

The distribution of other comorbidities shaped quite similarly between the 2 compared groups:

- In the CRA tool group, the other common risk factors were chronic lung disease (8337/248,862, 3.35%) and diabetes (5998/248,862, 2.41%).
- In the admission room group, the other common risk factors were diabetes (56/291, 19.24%), cancer (active neoplasms of all types, including of hematological origin; 30/291, 10.31%), and chronic lung disease (22/291, 7.56%).

A relatively high percentage of people reporting immunosuppression in the CRA tool group (weakened immune system; 14,708/248,862 [5.91%] of users) compared to the admission room group (6/291, 2.06%) suggests this risk factor might have been misinterpreted and misused despite the extensive description explaining the nature and examples of immunosuppression (available in [Multimedia Appendix 2](#)).

In general, admission room patients more often were burdened with comorbidities compared to CRA tool users. This can be explained by a higher average age of admission room patients compared to CRA tool users ([Table 3](#)).

Table 3. Distribution of comorbidities^a in the CRA^b tool and admission room groups.

Comorbidities	CRA tool			Admission room		
	COVID-19 positive (N=40,007), n (%)	COVID-19 negative (N=193,206), n (%)	P value ^c	COVID-19 positive (N=232), n (%)	COVID-19 negative (N=59), n (%)	P value ^d
Cardiovascular diseases	9346 (23.36)	26,296 (13.61)	<.001	125 (53.88)	13 (22.03)	<.001
Diabetes	1680 (4.20)	4012 (2.08)	<.001	51 (21.98)	5 (8.47)	.02
Current cancer	818 (2.04)	1517 (0.79)	<.001	28 (12.07)	2 (3.39)	.06
Diagnosed chronic lung disease	2461 (6.15)	5425 (2.81)	<.001	20 (8.62)	2 (3.39)	.27
History of chronic liver disease	1064 (2.66)	2140 (1.11)	<.001	7 (3.02)	1 (1.69)	.99
History of chronic kidney disease	967 (2.42)	1851 (0.96)	<.001	7 (3.02)	1 (1.69)	.99
Weakened immune system	4309 (10.77)	9629 (4.98)	<.001	5 (2.16)	0	0.59

^aOverall comorbidities: There were 20,645 comorbidities in COVID-19 positives and 50,870 comorbidities in COVID-19 negatives in the CRA tool group. There were 243 comorbidities in COVID-19 positives and 24 comorbidities in COVID-19 negatives in the admission room group.

^bCRA: COVID-19 risk assessment.

^cP values for CRA: test of proportions.

^dP values for the admission room: Fisher exact test.

Symptom Distribution

Overall, the most commonly reported symptoms differed between the CRA tool and the admission room groups. CRA interviews were dominated by mild symptoms, such as fatigue (61,544/248,862, 24.73%), cough (54,575/248,862, 21.93%), and headache (45,417/248,862, 18.25%). Meanwhile, the admission room patients presented with more serious symptoms, such as fever (175/291, 60.14%), cough (168/291, 57.73%), shortness of breath (114/291, 39.18%), and fatigue and muscle pain (59/291, 20.27% for both).

In the admission room group, the distribution of the most common symptoms among COVID-19-positive (232/291,

79.73%) and COVID-19-negative (59/291, 20.27%) patients was fairly similar: fever (n=146 [62.9%] of COVID-19 positives, n=29 [49.2%] of COVID-19 negatives), cough (n=138 [59.5%] of COVID-19 positives, n=30 [50.8%] of COVID-19 negatives), and shortness of breath (n=87 [37.5%] of COVID-19 positives, n=27 [45.8%] of COVID-19 negatives).

In contrast, the presentation of the COVID-19-suspected and COVID-19-nonsuspected individuals differed substantially. COVID-19-suspected users commonly reported symptoms such as fever, cough, and shortness of breath, while COVID-19-nonsuspected users commonly reported headache, cough, and fatigue. For details see [Table 4](#).

Table 4. Symptom and risk factor distribution of CRA^a tool users and admission room patients.

Symptom or risk factor	CRA tool		Admission room	
	COVID-19 positive (N=40,007), n (%)	COVID-19 negative (N=193,206), n (%)	COVID-19 positive (N=232), n (%)	COVID-19 negative (N=59), n (%)
Cough	25,062 (62.64)	29,521 (15.28)	138 (59.48)	30 (50.85)
Fever	23,123 (57.80)	20,292 (10.50)	146 (62.93)	29 (49.15)
Symptoms getting worse quickly	19,816 (49.53)	0	N/A ^b	N/A
Shortness of breath	15,157 (37.89)	12,717 (6.58)	87 (37.50)	27 (45.76)
Faster breathing	12,964 (32.40)	0	N/A	N/A
Fatigue	5987 (14.96)	52,630 (27.24)	40 (17.24)	19 (32.20)
Headache	4497 (11.24)	38,115 (19.73)	19 (8.19)	4 (6.78)
Sore throat	3975 (9.94)	35,645 (18.45)	17 (7.33)	9 (15.25)
Muscle pain	3351 (8.38)	27,015 (13.98)	42 (18.10)	17 (28.81)
Coughing up blood	2006 (5.01)	0	1 (0.43)	0
Chills	1906 (4.76)	13,740 (7.11)	2 (0.86)	2 (3.39)
Diarrhea	1242 (3.10)	14,109 (7.30)	35 (15.09)	4 (6.78)
Contact with infected person	1005 (2.51)	0	166 (71.55)	5 (8.47)
Nasal catarrh	954 (2.38)	6134 (3.17)	20 (8.62)	5 (8.47)
Loss of smell or taste	947 (2.37)	6034 (3.12)	39 (16.81)	2 (3.39)
Nausea	911 (2.28)	10,599 (5.49)	10 (4.31)	3 (5.08)
No contact with infected person	0	193,206 (100)	66 (28.45)	54 (91.53)

^aCRA: COVID-19 risk assessment.

^bN/A: not applicable.

Comparative Results

Fever and cough were the most commonly reported symptoms of COVID-19 in CRA tool users and admission room patients: fever occurred in 23,123/40,007 (57.80%) and 146/232 (62.93%) of the studied groups, respectively, while cough occurred in 25,062/40,007 (62.64%) and 138/232 (59.48%) of the studied groups, respectively. Pneumonia, characterized as the presence of fever, cough, and dyspnea, has been proven to be the most prevalent clinical presentation of COVID-19 in many studies [22-25].

Cardiovascular disease and diabetes occurred significantly more commonly in the COVID-19-positive than in the COVID-19-negative group both in the CRA tool (9346/40,007 [23.36%] vs 26,296/193,206 [13.61%] for cardiovascular disease, $P < .001$; 1680/40,007 [4.20%] vs 4012/40,007 [2.08%], $P < .001$ for diabetes) and in the admission room (125/232 [53.88%] vs 13/59 [22.03%] for cardiovascular disease; 51/232 [21.98%] vs 5/59 [8.47%], $P < .001$ for diabetes) group.

Anosmia or ageusia (2/59, 3.39%) occurred more frequently in the admission room group in COVID-19-positive than in COVID-19-negative patients. In the app, we did not observe a similar finding, probably due to the rapid cessation of the interview in high-triage scenarios.

Anosmia or ageusia occurred more frequently in mild than in severe COVID-19 in the CRA tool group (3849/40,007 [9.62%]

vs 40/40,007 [0.10%]). This is consistent with studies suggesting that olfactory and gustatory disturbances are among the most commonly reported symptoms in mild-to-moderate COVID-19 [26].

The average age of users of the COVID-19 self-assessment app was 37 years, whereas the average age of admission room patients was 58 years.

Fatigue, chills, nausea, and sore throat did not turn out to be diagnostically relevant for diagnosing COVID-19. In both CRA tool and admission room groups, they occurred more frequently in non-COVID-19 individuals.

Discussion

Principal Findings

The CRA tool ceased to be supported on August 16, 2021. As of now, most of the COVID-19 diagnostics are run by the Infermedica artificial intelligence (AI) engine [27], and the CRA tool is supported only in selected use cases (ie, the Polish Ministry of Health) [19].

The CRA tool, as it served as a means of screening and self-education, did not substitute for consultations in the admission room for symptomatic users. The tool could not confirm or exclude SARS-CoV-2 infection, as it cannot perform a laboratory examination. Hence, it does not substitute for physicians' interactions. However, our tool exercised the purpose

of remote triage. CRA did not overlook truly symptomatic cases; users with potentially worrisome symptoms, such as fever or shortness of breath, were identified and advised to obtain a consultation or schedule a COVID-19 test.

The compared groups—one that completed the online interview and one that reported to the hospital—differed in age distribution, the presence of risk factors, and probably the severity of symptoms reported. The difference between both groups impacted the results of the study, but it also showed some limitations of remote diagnostic tools, such as CRA—as patients potentially the most vulnerable to COVID-19 are also the least prevalent group accessing the internet for a health checkup. It is observed, however, that younger patients also suffer from COVID-19 infection, and with the next waves of pandemics, infections in young adults will become more prevalent [28]. This growing group of patients could have benefitted from remote triage assessment tools, such as CRA.

Taste and smell disorders occurred more commonly in the admission room group than in the CRA tool group (39/232 [16.81%] vs 947/40,007 [2.37%] for COVID-19-suspected individuals). In search of a possible explanation of this finding, we turned to the logic of WHO guidelines used in the CRA tool at that time. They did not distinguish smell and taste disorders as key diagnostic factors [15]. Once the importance of symptoms such as smell and taste disorders came to the attention of academics [29], WHO reflected these findings in the updated guidelines for suspecting COVID-19 infection (on August 7, 2020). WHO emphasized adjacent symptoms, such as diminished taste or smell, and reduced the significance of fever in suspecting COVID-19 infection. The newer versions of the CRA tool, not described in this paper, follow the guidelines, increasing their diagnostic importance.

It was not possible to assess the actual number of false-negative cases in the CRA tool due to a lack of data. However, we know that among the admission room records, 31 (13.78%) of 225 patients did not present with fever or dyspnea but still tested positive for COVID-19. These patients would have been classified as non-COVID-19 cases by the app.

Concomitant symptoms, such as fatigue, headache, and diarrhea, occurred infrequently in severe COVID-19-positive cases in the app. This may have been caused by the premature cessation of the interview for safety reasons.

The overall number of COVID-19-suspected cases in the CRA tool was 40,007 (16.08%) of 248,862 individual interviews. This number corresponds with the number of scheduled tests for novel coronavirus because in both these cases, we deal with the suspicion of COVID-19 based on presented history and symptoms. During a similar period, between May 11 and August 3, 2020, there were 17,864,205 tests for SARS-CoV-2 performed [30].

Limitations of the Study

Possible Misinterpretation of Red-Flag Questions

The outcomes of the self-assessment triage tool highlighted room for improvement with regard to phrasing questions in web apps for the common user. The “symptoms getting worse

quickly” red flag was meant to pinpoint a swiftly deteriorating user's general condition, which is a premise for hospitalization. However, a comparable number of confirmative and declined answers suggest that many of these answers could have been false positives. This answer might have been overly reported by the respondents, who may have misinterpreted its scope. In many cases, this occurrence may have led to the overtriage of urgent COVID-19 case recommendation (“Call the emergency number.”).

Bias of the Sample

As the tool was publicly available to everyone and no check-in or login was required, there is a possibility that some users did not present the symptoms they reported and used the tool only out of curiosity or for educational purposes. However, this bias is probably limited by the size of the group tested with the self-assessment tool.

More Detailed Screening in the Admission Room Sample

Screening in the admission room is always more exhaustive than in any self-assessment tool. There are a couple of contributing factors:

- Physical examinations cannot be substituted by any questions asked by the symptom checker.
- A general appearance provides valuable clinical information to experienced clinicians.
- There is a closed set of symptoms to choose from in the CRA tool.
- After detecting a potential red flag, the tool is designed to terminate the interview without inquiring about concomitant symptoms.

Conclusions

Comparing the symptomatology of users interacting with the CRA tool and those visiting the admission room revealed 2 major patient groups that could have benefitted from implementing the self-assessment app in preclinical triage settings.

The first group were patients with typical COVID-19 symptoms: cough and fever, sometimes accompanied by shortness of breath, tachypnea, fatigue, headache, and muscle pain. Some of these patients had additional comorbidities, such as diabetes or cardiovascular disease, that could have impacted the clinical course of COVID-19 [17]. The CRA tool could recognize patients with such symptoms as potentially COVID-19 positive and directed them to the proper care. The CRA tool was accurate in identifying patients at risk: every patient reporting a potential red-flag symptom, such as rapid symptom deterioration or acute dyspnea and tachypnea, was advised to seek immediate medical attention in the emergency room or was instructed to call the ambulance.

The other group were patients with no symptoms suggesting COVID-19 infection but still searching for answers as to whether they could be infected and what they should do. Oligosymptomatic and asymptomatic users, who constituted the majority of individuals interacting with the tool, were educated about their symptoms and advised to refer to the primary care in the case of symptom worsening. CRA has played

an educational role in advising on isolation precautions, organizing quarantine, and referring for further reading using evidence-based sources, such as WHO and the CDC.

It seems that these types of solutions may serve as health information hubs for oligosymptomatic individuals and means of remote triage for a vast audience. They possess the ability to identify patients at risk, providing them with next-step recommendations, as well as sieving out asymptomatic individuals, providing them with evidence-based education

materials. Such patients were the most prevalent (171,226 [68.80%] of the 248,862 CRA tool users).

As the study did not examine the intention of the user, it is uncertain what portion of such patients would visit a health care professional unnecessarily; further studies are required to assess the exact impact of online tools on reducing unnecessary visits. Still, as we observed oligosymptomatic patients visiting the hospital admission room, it can be assumed that some portion of such visits could be prevented by providing reassuring information to the patient through the online tool.

Acknowledgments

The Department of Infectious Diseases and Hepatology of the Wrocław Medical University provided funding for the statistical analysis. Infermedica provided funding for the submission fee.

Authors' Contributions

AN and JJ conceived and presented the idea. JJ collected and analyzed the data regarding the self-assessment tool. AN collected and analyzed the data regarding the Infectious Diseases Admission Room. AN and JJ wrote the first version of the manuscript. ASP and KS supervised the work, provided new ideas regarding Discussion and Conclusion sections, and helped with result interpretation. All authors discussed the results and contributed to the final manuscript.

Conflicts of Interest

AN and JJ are affiliated professionally with Infermedica as medical consultants. They have contributed to creating the COVID-19 risk assessment (CRA) tool by outlining and adapting the medical foundation based on World Health Organization (WHO) guidelines for COVID-19 surveillance. They did not receive any compensation for the study.

AN and ASP are physicians at the First Infectious Diseases Department in J. Gromkowski Hospital in Wrocław. ASP is an assistant at the Department of Infectious Diseases and Hepatology of the Wrocław Medical University.

KS is physician and the head of the First Department of Infectious Diseases in J. Gromkowski Hospital in Wrocław, as well as the head of the Department of Infectious Diseases and Hepatology of the Wrocław Medical University.

Multimedia Appendix 1

Extensive screen description and decision tree logic.

[[DOCX File , 1795 KB](#) - [humanfactors_v9i2e34134_app1.docx](#)]

Multimedia Appendix 2

Distribution of symptoms and comorbidities in the CRA tool, displayed by the triage outcome. CRA: COVID-19 risk assessment.

[[XLSX File \(Microsoft Excel File\), 13 KB](#) - [humanfactors_v9i2e34134_app2.xlsx](#)]

References

1. Blumenthal D, Fowler EJ, Abrams M, Collins SR. Covid-19: implications for the health care system. *N Engl J Med* 2020 Oct 08;383(15):1483-1488 [[FREE Full text](#)] [doi: [10.1056/nejmsb2021088](https://doi.org/10.1056/nejmsb2021088)]
2. Moynihan R, Sanders S, Michaleff ZA, Scott AM, Clark J, To EJ, et al. Impact of COVID-19 pandemic on utilisation of healthcare services: a systematic review. *BMJ Open* 2021 Mar 16;11(3):e045343 [[FREE Full text](#)] [doi: [10.1136/bmjopen-2020-045343](https://doi.org/10.1136/bmjopen-2020-045343)] [Medline: [33727273](https://pubmed.ncbi.nlm.nih.gov/33727273/)]
3. All-in-One SEO Toolset. URL: <https://ahrefs.com/> [accessed 2022-03-17]
4. Alexander GC, Tajanlangit M, Heyward J, Mansour O, Qato DM, Stafford RS. Use and content of primary care office-based vs telemedicine care visits during the COVID-19 pandemic in the US. *JAMA Netw Open* 2020 Oct 01;3(10):e2021476 [[FREE Full text](#)] [doi: [10.1001/jamanetworkopen.2020.21476](https://doi.org/10.1001/jamanetworkopen.2020.21476)] [Medline: [33006622](https://pubmed.ncbi.nlm.nih.gov/33006622/)]
5. Kaboli PJ, Augustine MR, Rose DE, Chawla N, Bouchard M, Hebert P. Call center remote triage by nurse practitioners was associated with fewer subsequent face-to-face healthcare visits. *J Gen Intern Med* 2021 Aug;36(8):2315-2322 [[FREE Full text](#)] [doi: [10.1007/s11606-020-06536-0](https://doi.org/10.1007/s11606-020-06536-0)] [Medline: [33501532](https://pubmed.ncbi.nlm.nih.gov/33501532/)]
6. Judson T, Odisho AY, Neinstein AB, Chao J, Williams A, Miller C, et al. Rapid design and implementation of an integrated patient self-triage and self-scheduling tool for COVID-19. *J Am Med Inform Assoc* 2020 Jun 01;27(6):860-866 [[FREE Full text](#)] [doi: [10.1093/jamia/ocaa051](https://doi.org/10.1093/jamia/ocaa051)] [Medline: [32267928](https://pubmed.ncbi.nlm.nih.gov/32267928/)]

7. Stemler J, Cornely OA, Noack-Schönborn T, Fohrholz C, Schumacher S, Poluschkin L, et al. Web-based, rapid and contactless management of ambulatory patients for SARS-CoV-2-testing. *BMC Infect Dis* 2021 Jun 07;21(1):535 [FREE Full text] [doi: [10.1186/s12879-021-06249-7](https://doi.org/10.1186/s12879-021-06249-7)] [Medline: [34098882](https://pubmed.ncbi.nlm.nih.gov/34098882/)]
8. Hautz W, Exadaktylos A, Sauter TC. Online forward triage during the COVID-19 outbreak. *Emerg Med J* 2021 Feb;38(2):106-108 [FREE Full text] [doi: [10.1136/emmermed-2020-209792](https://doi.org/10.1136/emmermed-2020-209792)] [Medline: [33310732](https://pubmed.ncbi.nlm.nih.gov/33310732/)]
9. Lian W, Wen L, Zhou Q, Zhu W, Duan W, Xiao X, et al. Digital health technologies respond to the COVID-19 pandemic in a tertiary hospital in China: development and usability study. *J Med Internet Res* 2020 Nov 24;22(11):e24505 [FREE Full text] [doi: [10.2196/24505](https://doi.org/10.2196/24505)] [Medline: [33141679](https://pubmed.ncbi.nlm.nih.gov/33141679/)]
10. Monaghesh E, Hajizadeh A. The role of telehealth during COVID-19 outbreak: a systematic review based on current evidence. *BMC Public Health* 2020 Aug 01;20(1):1193 [FREE Full text] [doi: [10.1186/s12889-020-09301-4](https://doi.org/10.1186/s12889-020-09301-4)] [Medline: [32738884](https://pubmed.ncbi.nlm.nih.gov/32738884/)]
11. Mehring W, Poksay A, Kriege J, Prasannappa R, Wang MD, Hendel C, et al. Initial experience with a COVID-19 web-based patient self-assessment tool. *J Gen Intern Med* 2020 Sep;35(9):2821-2822 [FREE Full text] [doi: [10.1007/s11606-020-05893-0](https://doi.org/10.1007/s11606-020-05893-0)] [Medline: [32542495](https://pubmed.ncbi.nlm.nih.gov/32542495/)]
12. Schrage J, Schuler K, Isakov AP, Wright DW, Yaffee AQ, Jacobson KL, et al. Development and usability testing of a web-based COVID-19 self-triage platform. *West J Emerg Med* 2020 Aug 19;21(5):1054-1058 [FREE Full text] [doi: [10.5811/westjem.2020.7.48217](https://doi.org/10.5811/westjem.2020.7.48217)] [Medline: [32970554](https://pubmed.ncbi.nlm.nih.gov/32970554/)]
13. Lunn PD, Timmons S, Julianne H, Belton CA, Barjaková M, Lavin C, et al. Using decision aids to support self-isolation during the COVID-19 pandemic. *Psychol Health* 2021 Feb 19;36(2):195-213. [doi: [10.1080/08870446.2020.1849701](https://doi.org/10.1080/08870446.2020.1849701)] [Medline: [33210950](https://pubmed.ncbi.nlm.nih.gov/33210950/)]
14. Komisja Bioetyczna UMP (Bioethical Committee of Poznan Medical University). Badania Naukowe Niesponsorowane (Non-funded studies). URL: http://www.bioetyka.ump.edu.pl/BADANIA_NAUKOWE_NIESPONSOROWANE.html [accessed 2022-03-13]
15. World Health Organization. Public Health Surveillance for COVID-19: Interim Guidance. URL: <https://www.who.int/publications/i/item/WHO-2019-nCoV-SurveillanceGuidance-2022.1> [accessed 2022-03-17]
16. World Health Organization. Situation Reports. URL: <https://www.who.int/emergencies/diseases/novel-coronavirus-2019/situation-reports> [accessed 2022-03-17]
17. Grasselli G, Greco M, Zanella A, Albano G, Antonelli M, Bellani G, COVID-19 Lombardy ICU Network. Risk factors associated with mortality among patients with COVID-19 in intensive care units in Lombardy, Italy. *JAMA Intern Med* 2020 Oct 01;180(10):1345-1355 Erratum in: *JAMA Intern Med*. 2021 Jul 1;181(7):1021 [FREE Full text] [doi: [10.1001/jamainternmed.2020.3539](https://doi.org/10.1001/jamainternmed.2020.3539)] [Medline: [32667669](https://pubmed.ncbi.nlm.nih.gov/32667669/)]
18. Fine M, Auble TE, Yealy DM, Hanusa BH, Weissfeld LA, Singer DE, et al. A prediction rule to identify low-risk patients with community-acquired pneumonia. *N Engl J Med* 1997 Jan 23;336(4):243-250 [FREE Full text] [doi: [10.1056/nejm199701233360402](https://doi.org/10.1056/nejm199701233360402)]
19. Polish Ministry of Health. Sprawdź, czy masz objawy COVID-19. URL: <https://pacjent.gov.pl/koronawirus/sprawdz-objawy> [accessed 2022-03-17]
20. Infermedica. COVID-19 Risk Assessment Tool. URL: <https://covid19.infermedica.com/> [accessed 2022-03-17]
21. Infermedica. COVID-19 Risk Assessment: 3 Weeks from Conception to Launch. URL: <https://infermedica.com/case-studies/covid19> [accessed 2022-03-17]
22. Huang C, Wang Y, Li X, Ren L, Zhao J, Hu Y, et al. Clinical features of patients infected with 2019 novel coronavirus in Wuhan, China. *Lancet* 2020 Feb;395(10223):497-506 [FREE Full text] [doi: [10.1016/s0140-6736\(20\)30183-5](https://doi.org/10.1016/s0140-6736(20)30183-5)]
23. Chen N, Zhou M, Dong X, Qu J, Gong F, Han Y, et al. Epidemiological and clinical characteristics of 99 cases of 2019 novel coronavirus pneumonia in Wuhan, China: a descriptive study. *Lancet* 2020 Feb;395(10223):507-513 [FREE Full text] [doi: [10.1016/s0140-6736\(20\)30211-7](https://doi.org/10.1016/s0140-6736(20)30211-7)]
24. Wang D, Hu B, Hu C, Zhu F, Liu X, Zhang J, et al. Clinical characteristics of 138 hospitalized patients with 2019 novel coronavirus-infected pneumonia in Wuhan, China. *JAMA* 2020 Mar 17;323(11):1061-1069 [FREE Full text] [doi: [10.1001/jama.2020.1585](https://doi.org/10.1001/jama.2020.1585)] [Medline: [32031570](https://pubmed.ncbi.nlm.nih.gov/32031570/)]
25. Guan W, Ni Z, Hu Y, Liang W, Ou C, He J, et al. Clinical characteristics of coronavirus disease 2019 in China. *N Engl J Med* 2020 Apr 30;382(18):1708-1720 [FREE Full text] [doi: [10.1056/nejmoa2002032](https://doi.org/10.1056/nejmoa2002032)]
26. Lechien JR, Chiesa-Estomba CM, De Siati DR, Horoi M, Le Bon SD, Rodriguez A, et al. Olfactory and gustatory dysfunctions as a clinical presentation of mild-to-moderate forms of the coronavirus disease (COVID-19): a multicenter European study. *Eur Arch Otorhinolaryngol* 2020 Aug 06;277(8):2251-2261 [FREE Full text] [doi: [10.1007/s00405-020-05965-1](https://doi.org/10.1007/s00405-020-05965-1)] [Medline: [32253535](https://pubmed.ncbi.nlm.nih.gov/32253535/)]
27. Symptomate. What Concerns You about Your Health Today?. URL: <https://symptomate.com/> [accessed 2022-03-17]
28. Twohig KA, Nyberg T, Zaidi A, Thelwall S, Sinnathamby MA, Aliabadi S, COVID-19 Genomics UK (COG-UK) consortium. Hospital admission and emergency care attendance risk for SARS-CoV-2 delta (B.1.617.2) compared with alpha (B.1.1.7) variants of concern: a cohort study. *Lancet Infect Dis* 2022 Jan;22(1):35-42. [doi: [10.1016/s1473-3099\(21\)00475-8](https://doi.org/10.1016/s1473-3099(21)00475-8)]

29. Tong JY, Wong A, Zhu D, Fastenberg JH, Tham T. The prevalence of olfactory and gustatory dysfunction in COVID-19 patients: a systematic review and meta-analysis. *Otolaryngol Head Neck Surg* 2020 Jul 05;163(1):3-11 [FREE Full text] [doi: [10.1177/0194599820926473](https://doi.org/10.1177/0194599820926473)] [Medline: [32369429](https://pubmed.ncbi.nlm.nih.gov/32369429/)]
30. Polish Ministry of Health. A Number of Performed COVID-19 Tests. URL: <https://www.gov.pl/attachment/6f9f0ab7-20b9-4681-a866-38d8f599d034> [accessed 2022-03-17]

Abbreviations

API: Application Program Interface
CDC: Centers for Disease Control and Prevention
CRA: COVID-19 risk assessment
WHO: World Health Organization

Edited by A Kushniruk; submitted 07.10.21; peer-reviewed by M Schmieding, K Patel, R Menhas; comments to author 31.10.21; revised version received 30.12.21; accepted 13.02.22; published 04.04.22.

Please cite as:

Nowicka A, Jaszczak J, Szymanek Pasternak A, Simon K

Application of a Web-based Self-assessment Triage Tool During the COVID-19 Pandemic: Descriptive Study

JMIR Hum Factors 2022;9(2):e34134

URL: <https://humanfactors.jmir.org/2022/2/e34134>

doi: [10.2196/34134](https://doi.org/10.2196/34134)

PMID: [35168192](https://pubmed.ncbi.nlm.nih.gov/35168192/)

©Anna Nowicka, Jakub Jaszczak, Anna Szymanek Pasternak, Krzysztof Simon. Originally published in JMIR Human Factors (<https://humanfactors.jmir.org>), 04.04.2022. This is an open-access article distributed under the terms of the Creative Commons Attribution License (<https://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Human Factors, is properly cited. The complete bibliographic information, a link to the original publication on <https://humanfactors.jmir.org>, as well as this copyright and license information must be included.

Original Paper

Patient Portal Messaging for Asynchronous Virtual Care During the COVID-19 Pandemic: Retrospective Analysis

Ming Huang¹, PhD; Aditya Khurana², MD; George Mastorakos², MD; Andrew Wen¹, MS; Huan He¹, PhD; Liwei Wang¹, PhD; Sijia Liu¹, PhD; Yanshan Wang¹, PhD; Nansu Zong¹, PhD; Julie Prigge³, MA; Brian Costello³, MD; Nilay Shah¹, PhD; Henry Ting⁴, MD; Jungwei Fan¹, PhD; Christi Patten^{5,6}, PhD; Hongfang Liu¹, PhD

¹Department of Artificial Intelligence and Informatics, Mayo Clinic, Rochester, MN, United States

²Mayo Clinic Alix School of Medicine, Mayo Clinic, Scottsdale, AZ, United States

³Center for Connected Care, Mayo Clinic, Rochester, MN, United States

⁴Department of Cardiovascular Medicine, Mayo Clinic, Rochester, MN, United States

⁵Center for Clinical and Translational Science, Mayo Clinic, Rochester, MN, United States

⁶Department of Psychiatry and Psychology, Mayo Clinic, Rochester, MN, United States

Corresponding Author:

Hongfang Liu, PhD

Department of Artificial Intelligence and Informatics

Mayo Clinic

200 1st Street SW

Rochester, MN, 55905

United States

Phone: 1 507 293 0057

Email: Liu.Hongfang@mayo.edu

Abstract

Background: During the COVID-19 pandemic, patient portals and their message platforms allowed remote access to health care. Utilization patterns in patient messaging during the COVID-19 crisis have not been studied thoroughly. In this work, we propose characterizing patients and their use of asynchronous virtual care for COVID-19 via a retrospective analysis of patient portal messages.

Objective: This study aimed to perform a retrospective analysis of portal messages to probe asynchronous patient responses to the COVID-19 crisis.

Methods: We collected over 2 million patient-generated messages (PGMs) at Mayo Clinic during February 1 to August 31, 2020. We analyzed descriptive statistics on PGMs related to COVID-19 and incorporated patients' sociodemographic factors into the analysis. We analyzed the PGMs on COVID-19 in terms of COVID-19-related care (eg, COVID-19 symptom self-assessment and COVID-19 tests and results) and other health issues (eg, appointment cancellation, anxiety, and depression).

Results: The majority of PGMs on COVID-19 pertained to COVID-19 symptom self-assessment (42.50%) and COVID-19 tests and results (30.84%). The PGMs related to COVID-19 symptom self-assessment and COVID-19 test results had dynamic patterns and peaks similar to the newly confirmed cases in the United States and in Minnesota. The trend of PGMs related to COVID-19 care plans paralleled trends in newly hospitalized cases and deaths. After an initial peak in March, the PGMs on issues such as appointment cancellations and anxiety regarding COVID-19 displayed a declining trend. The majority of message senders were 30-64 years old, married, female, White, or urban residents. This majority was an even higher proportion among patients who sent portal messages on COVID-19.

Conclusions: During the COVID-19 pandemic, patients increased portal messaging utilization to address health care issues about COVID-19 (in particular, symptom self-assessment and tests and results). Trends in message usage closely followed national trends in new cases and hospitalizations. There is a wide disparity for minority and rural populations in the use of PGMs for addressing the COVID-19 crisis.

(*JMIR Hum Factors* 2022;9(2):e35187) doi:[10.2196/35187](https://doi.org/10.2196/35187)

KEYWORDS

patient portal; patient portal message; asynchronous communication; COVID-19; utilization; digital health; healthcare; health care; remote healthcare; virtual care; pandemic

Introduction

The COVID-19 pandemic accelerated the adoption of digital and virtual patient care technology as sustainable and scalable parts of health systems. This includes the use of video, audio, and even Health Insurance Portability and Accountability Act–secure portals as a means for patients to remain connected with their providers [1]. Compared with synchronous virtual care such as phone and video visits [2,3], patient portals as secure web-based platforms allow patients to conveniently access information from their electronic health records and asynchronously interact with their providers [4]. Patient portals are becoming increasingly common and give patients unlimited access to their health information (eg, clinical notes, test results, medications, and discharge summaries) from anywhere through an internet connection [5]. A study shows that over 90% of health care organizations (eg, Veterans Administration and Kaiser Permanente) had provided patient portal services to their patients [6]. Convenient access and management of personal health information have been shown to improve patients' self-management of diseases by increasing awareness of disease knowledge, status, and progress [7]. Additionally, patient portals provide a significant function of portal messaging for asynchronous communication between patients and their providers or care teams on a wide spectrum of tasks such as appointment requests, virtual visits, care management, or mental health issues [8-10].

Millions of nonurgent and non-COVID-19 medical encounters were postponed or cancelled by patients and health systems to reduce the risk of COVID-19 infection during in-person visits and prevent virus spread [11-14]. For continued health care access, most clinic visits had transitioned to web-based platforms for health care access, including COVID-19 diagnosis and treatment [2,15]. Through patient portals, patients could receive educational information on COVID-19 preventive care measures or use web-based triage forms (e-visit) for COVID-19 symptom assessment by an advanced practice provider. If a COVID-19 diagnostic test was recommended (and completed), patients could send and receive portal messages related to their COVID-19 diagnostic tests and test results. Even with a positive test result, patients could communicate with their providers about their COVID-19 symptoms through remote patient monitoring. In addition, the patient portal also provided a COVID-19 symptom self-assessment that patients could use interactively.

In the early stages of the global response, the public health strategy involved isolation for those infected or at risk, reducing social contact to slow the spread, and masking and hand washing to reduce infection risk. However, this unintentionally led to increased feelings of loneliness, reduced access social support, and worsening stress, anxiety, and depressive symptoms [16-18]. Studies indicate that the COVID-19 crisis and resulting economic and social lockdown and isolation had negatively impacted patients' mental health [19-21]. Patients may be using

patient portals to interact with their providers about their mental health conditions to seek support.

Studies on the use of telehealth and patient portal technologies have recently increased during the COVID-19 pandemic [2,15,22]. For example, Patel et al [23] implemented telehealth capabilities for COVID-19 care within their pediatric patient portal and found that weekly telehealth visits subsequently increased 200-fold for children and 90-fold for adolescents. Khairat et al [24] analyzed the use of pediatric tele-urgent care visits via a patient portal at a southeastern health care center and revealed that the use of tele-urgent care visits for pediatric care doubled during the COVID-19 crisis. Portz et al [25] observed a large increase of patient portal utilization for advance care planning. However, very few research studies examined the messaging component of the portal, an important function of patient portals for asynchronous communication between patients and providers, specifically for COVID-19–related care and issues [26]. Thus, in this study, we assessed portal messages associated with COVID-19 generated by patients from February 1 to August 31, 2020, at Mayo Clinic, a large multispecialty academic health system. We summarized reasons for patient utilization of portal messaging for COVID-19–related care such as diagnosis, testing, treatment, scheduling, and mental health support [27]. During this time period, vaccinations for COVID-19 were not available; hence, this topic was not analyzed. In addition, we analyzed patient user demographics with respect to their personal and social factors such as age, gender, race, and geographic location. These findings can provide insight into how patients interacted with the asynchronous portion of the patient portal during the COVID-19 pandemic.

Methods

Data Collection and Preprocessing

Mayo Clinic is a large multispecialty academic medical center focused on integrated patient care, education, and research. Mayo Clinic has three main medical sites in Minnesota, Florida, and Arizona and Mayo Clinic Health System (MCHS). MCHS is as a network of community-based medical services and consisted of more than 40 hospitals and clinics in Minnesota, Iowa, and Wisconsin in 2021. Mayo Clinic's patient portal (Patient Online Services) has been operational since 2010 [28]. We collected over 2 million portal messages generated by patients from the Epic Clarity database between February 1 and August 31, 2020. We filtered the patient-generated messages (PGMs) associated with COVID-19 using relevant keywords (eg, "COVID-19," "Pandemic," "Coronavirus," "SARS-CoV-2," and "2019-nCoV") and their synonyms and morphological variations (see Table S1 in the [Multimedia Appendix 1](#)). We excluded the PGMs with empty message bodies and the PGMs requested by providers such as messages for preappointment COVID-19 screening and postdischarge COVID-19 symptom checks. We then identified 207,299 portal messages on

COVID-19 generated by 102,470 patients from the Epic Clarity database. In addition, the patient portal provided an anonymous COVID-19 self-checker for patients' self-assessment of COVID-19 symptoms on March 22, 2020. We collected 153,224 PGMs on COVID-19 symptom self-assessment during March 22 to August 31, 2020. Thus, a total of 360,523 PGMs were used for sequential analysis.

Data Analysis

Sociodemographic Characteristics of Patients

We analyzed the distribution of unique patients by age, gender, marriage, ethnicity, race, language, and residence. We excluded anonymous patients who sent messages for COVID-19 symptom self-assessment in the demographic analysis. We conducted a subanalysis comparing three different cohorts within our sample: patients who sent messages related to COVID-19 only (COVID-19 message senders), any patients who sent messages related to any topic (general message senders), and all patients who were active on the portal regardless of whether they composed messages (general patients). Statistical analysis involved chi-square goodness-of-fit tests.

PGMs Related to COVID-19

We calculated the daily numbers of total PGMs related to COVID-19 between February 1 and August 31, 2020. The daily numbers would exhibit a week periodicity (typically with a maximum on Monday and a minimum on weekends). Because of this, we calculated their weekly smoothing averages (WSAs). The WSAs displayed a reduction around the holidays (Memorial Day on May 25, 2020, and Independence Day on July 4, 2020); hence, holidays were excluded from the analysis. The daily numbers and WSAs of the PGMs on COVID-19 can approximate the overall utilization of portal messages by the patients for addressing the COVID-19 crisis over time.

Messages for COVID-19–Related Care and Other Health Care Issues

We analyzed the PGMs used for assessing COVID-19 symptoms and discussing COVID-19 care plans to understand the message utilization for COVID-19 diagnosis and treatment. We filtered the PGMs on COVID-19 symptom assessment by searching the relevant phrase, "COVID-19 (Coronavirus) Symptom Assessment," as well as relevant keywords such as "test" and "result" for diagnostic tests and results and "care plan," "monitoring," and "interactive care" for care plans (see Table S1 in the [Multimedia Appendix 1](#)). We then calculated the daily numbers and WSAs of these PGMs.

In addition, we examined other health care issues caused by COVID-19 reported in the portal messages to understand the impacts of the COVID-19 pandemic on health services and patients. We calculated the number of PGMs explicitly mentioning the phrase "due to COVID-19" and its synonyms to examine patient-reported health care issues caused by the

COVID-19 pandemic. We also computed the number of PGMs on COVID-19, which discussed rescheduling or cancelling appointments, mental health, and suicidal ideation using relevant keywords (eg, "cancel" and "reschedule" for appointments, "anxiety" and "depressed" for mental health, and "suicide" for suicidal ideation) and their synonyms to quantify the impact of the COVID-19 pandemic on health services and patients (see Table S1 in the [Multimedia Appendix 1](#)).

Evaluation

We recruited 2 medical students for annotating the binary code for each studied topic in portal messages: whether a portal message is linked to COVID-19, COVID-19–related care, or other health care issues due to COVID-19. We randomly sampled 1800 portal messages for annotation, and the results are shown in Table S2 in [Multimedia Appendix 1](#). The first 100 portal messages (10 each topic) were sampled and labeled by both annotators and their overall interagreement score is 0.91. After that, the two annotators worked together to complete the rest of the annotation. More specifically, we randomly sampled 100 portal messages on COVID-19 and 900 portal messages not related to COVID-19, which were identified by using the keyword searching for labeling. The precision, recall, and F1-score of the COVID-19 keyword search was over 0.99. We randomly sampled 100 portal messages on COVID-19 linked to each topic such as "isolation." The F1-scores range from 63.1% to 94.9%, except for "symptom assessment" (100%). The portal message on symptom assessment has a special phrase, "E-Visit Submission: COVID-19 (Coronavirus) Symptom Assessment," to accurately filter.

Ethics Approval

No patients were exposed to any intervention. We used the data from the Mayo Clinic Unified Data Platform for analysis. The study was approved by the Mayo Clinic institutional review board (19-002211).

Results

Patient Characteristics

[Table 1](#) lists the demographic distribution of patient populations from Epic Clarity database by age, gender, marriage, ethnicity, race, language, and residence. The results of chi-square goodness-of-fit tests among the three patient role groups (COVID-19 message senders, general message senders, and general patients) were listed in [Table S2](#) in [Multimedia Appendix 1](#).

We found that both COVID-19 and general message senders had a significantly different distribution compared to all patients active on the portal ($P<.001$). The distribution of the COVID-19 message senders was also significantly different from that of general message senders in terms of age, gender, marriage, race, and residence ($P<.001$).

Table 1. Demographic distribution of patients: COVID-19 message senders, general message senders, and general patients.

Patient demographics	COVID-19 message senders (N=102,470), %	General message senders (N=384,922), %	General patients (N=1,055,319), %
Age			
<18	7.58	8.25	10.67
18-29	9.80	9.75	10.22
30-39	13.14	11.81	10.36
40-49	13.84	12.94	11.45
50-64	29.11	28.05	24.83
≥65	26.53	29.20	32.47
Gender			
Female	60.97	58.54	54.37
Male	39.03	41.46	45.63
Marital status			
Married or has a life partner	63.87	63.22	56.06
Not married or legally separated	36.13	36.78	43.94
Ethnicity			
Non-Hispanic or Latino	95.84	95.82	95.28
Hispanic or Latino	4.16	4.18	4.72
Race			
White	92.73	92.40	90.84
Asian	2.44	2.42	2.28
Black or African American	2.11	2.44	3.40
American Indian or Alaska Native	0.37	0.37	0.41
Native Hawaiian or Pacific Islander	0.09	0.10	0.12
Other	2.26	2.28	2.94
Language			
English	99.09	99.06	97.56
Arabic	0.12	0.14	0.25
Spanish	0.34	0.37	1.03
Other	0.45	0.44	1.16
Residential area			
Urban	74.77	70.41	61.02
Rural	25.23	29.59	38.98

More than half (>55%) of patients were in the age groups of 50-64 years and ≥65 years. The proportion of patients in the age ranges of 30-39 years, 40-49 years, and 50-59 years was observed to have increased when looking at general portal users to general message senders to COVID-19-specific message senders. Meanwhile, the proportion of message senders in the age groups of <18 years and ≥65 years was the lowest in the COVID-19 message sender cohort.

More than half (>54%) of the patients were female and were married or had a life partner. The proportion of female patients (61% vs 54%) and married patients (64% vs 56%) also was highest in the COVID-19 message sender cohort when compared

to the entire active portal user cohort. More than 90% of patients were of non-Hispanic or Latino ethnicity, White race, and spoke English. This proportion was also highest in the COVID-19 message sender cohort and lowest in the general portal cohort.

At least 61% of all patients assessed lived in the urban area. The percentage of urban patients increased to 70% among general message senders and 75% among COVID-19 message senders.

PGMs Related to COVID-19

We illustrated the daily numbers and WSAs of PGMs on COVID-19 in [Figure 1](#). The WSA of PGMs on COVID-19 started to increase at the end of February and quickly peaked

at 3303.43 messages per week on March 30. Sequentially, the WSA of PGMs on COVID-19 decreased to a local minimum of 1750.29 messages per week on June 5. The WSA of PGMs on COVID-19 reached a local maximum of 2624.57 messages per week on July 12. This July peak was 79.45% of the March

peak’s volume. The WSA of PGMs on COVID-19 displayed a consistently declining trend over time after this point. [Table 2](#) lists the numbers and proportions of PGMs for COVID-19–related care and other health care issues.

Figure 1. Daily numbers and weekly smoothing averages (WSAs) of patient-generated messages (PGMs) related to COVID-19.

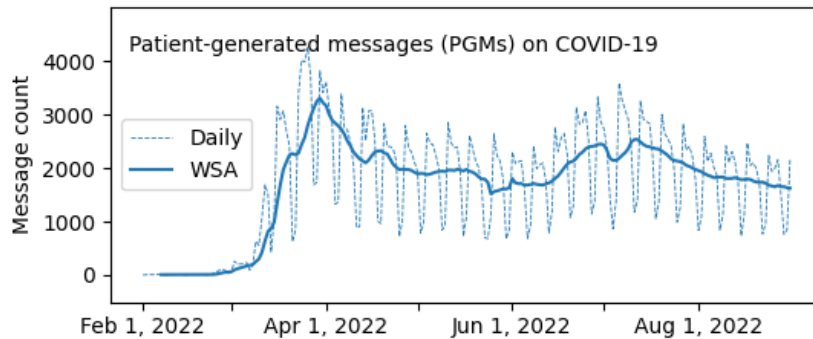


Table 2. Patient-generated messages (PGMs) related to COVID-19–related care and other health care issues caused by COVID-19.

Category	PGMs on COVID-19 (N=360,523), n (%)
COVID-19–related care	
Self-checker	153,224 (42.50)
e-Visit	4619 (1.28)
Tests and results	111,183 (30.84)
Care plan	3844 (1.07)
Other issues caused by the COVID-19 pandemic	
General issues	13,333 (3.70)
Postponement	26,924 (7.47)
Cancellation	19,000 (5.27)
Anxiety	21,413 (5.94)
Depression	3673 (1.02)
Suicidal ideation	288 (0.08)

Messages for COVID-19–Related Care and Other Health Care Issues

[Figure 2](#) depicts the daily numbers and WSAs of PGMs for COVID-19 symptom assessment via the self-checker ([Figure 2A](#)), COVID-19 symptom assessment via e-visits ([Figure 2B](#)), discussing the COVID-19 diagnostic tests and results ([Figure 2C](#)), and the care plan ([Figure 2D](#)). The top message concepts for COVID-19–related care were COVID-19 symptom assessment via the self-checker (42.50%) and COVID-19 tests and results (30.84%). The percentage of PGMs on COVID-19 symptom assessment associated with e-visits or the COVID-19 care plan was approximately 1%. Owing to the low use of COVID-19 symptom assessment via e-visits, the patient portal stopped the e-visit service for COVID-19 symptom assessment on August 1, 2020.

Similar to the total PGMs on COVID-19 in [Figure 1](#), the PGMs on COVID-19 symptom assessment via self-checker and COVID-19 tests and results had analogous dynamic patterns: two peaks in late March or early April and late June or early

July. These fluctuations were consistent with the surge in newly confirmed COVID-19 cases in the United States and in Minnesota (see [Figure S1](#) in [Multimedia Appendix 1](#)) [11]. The PGMs for COVID-19 symptom assessment via e-visits had two similar peaks and surges before the termination of the service. The trend of PGMs related to COVID-19 care plans paralleled trends in newly hospitalized cases and deaths in the United States and in Minnesota (see [Figure S1](#) in the [Multimedia Appendix 1](#)) [11].

[Table 2](#) shows that the top message usage for other health care issues caused by COVID-19 was related to appointment postponement (7.47%), anxiety (5.94%), and appointment cancellation (5.27%). Among the studied mental health issues, the number of relevant PGMs decreased as the severity of health issues increased: anxiety (5.94%), depression (1.02%), and suicidal ideation (0.08%). [Figure 3](#) depicts the numbers of PGMs related to COVID-19, which mentioned general issues ([Figure 3A](#)), appointment postponement ([Figure 3B](#)), appointment cancellation ([Figure 3C](#)), anxiety ([Figure 3D](#)), depression ([Figure 3E](#)), and suicidal ideation ([Figure 3F](#)). The curves in

Figure 3 show a similar trend over time: the number of PGMs on COVID-19 started to increase in early March and quickly peaked within 2-3 weeks. However, after March 30, the number of PGMs on COVID-19 constantly decreased, although smaller

upward fluctuations occurred from July to August in some curves; for example, such as PGMs for postponement, anxiety, and depression. These fluctuations paralleled those observed in PGM use for COVID-19 care plans, as shown in Figure 2D.

Figure 2. Daily numbers and weekly smoothing averages (WSAs) of patient-generated messages (PGMs) regarding COVID-19–related care (diagnosis and treatment): (A) COVID-19 symptom assessment via self-checker, (B) COVID-19 symptom assessment by providers via e-visits, (C) discussions regarding COVID-19 tests and results, and (D) care plans.

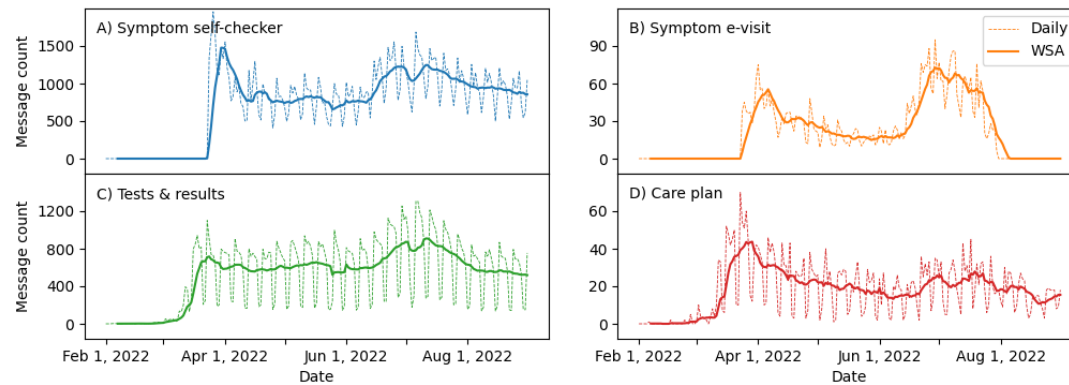
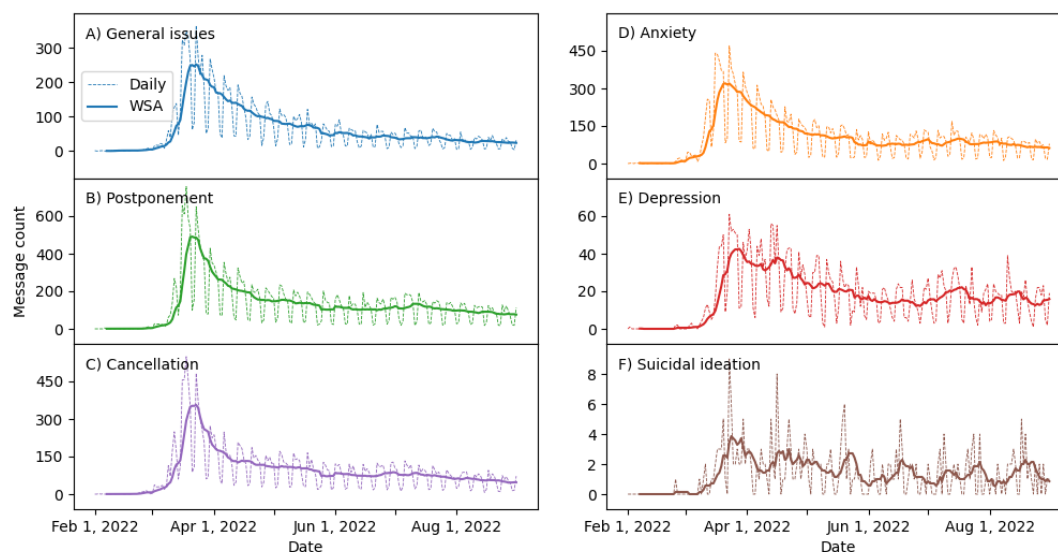


Figure 3. Daily numbers and weekly smoothing averages (WSAs) of patient-generated messages (PGMs) regarding COVID-19–related other health care issues: (A) general issues due to COVID-19, (B) postponement, (C) cancellation, (D) anxiety, (E) depression, and (F) suicidal ideation.



Discussion

Principal Findings

The COVID-19 pandemic and subsequent public health mitigation strategies, including stay-at-home orders and business restrictions, substantially impacted delivery of health care services. As the COVID-19 pandemic progressed in the United States, and specifically in Minnesota, newly confirmed cases had two peaks during February 1 and August 31, 2020, owing to the initial outbreak and late termination of stay-at-home orders. We observed similar dynamic patterns in PGMs on COVID-19, particularly, COVID-19 diagnosis and treatment, suggesting that patients actively used the portal messaging for addressing their concerns regarding the COVID-19 crisis [23].

Another previous study [26] was consistent with our findings, as they analyzed messages in an ambulatory practice network and determined that their inbox message usage patterns were also consistent with national trends. Patients sent portal messages mainly for COVID-19 symptom self-assessment and discussing COVID-19 tests and results. It appeared that patients preferred symptom self-assessment to e-visits for symptom assessment given the utilization rates. Thus, analyzing PGMs on COVID-19 symptom assessment via self-checker before diagnostic testing could serve as a timely surveillance of COVID-19. Prior work by Denis et al [29] utilized this association when developing a self-assessment web-based app to assess trends of the COVID-19 pandemic in France. We also determined that the PGMs related to COVID-19 care plans

followed trends in newly hospitalized cases and deaths. The second relative maximum of PGMs on the COVID-19 care plan after May suggests a decline of COVID-19 risk during that time period.

Our findings also indicated that patients used the portal to report feelings of anxiety and depression about their existing medical conditions and potential contagious risks due to COVID-19 and seek support from their providers. Similar mental health concerns increased in the general population, according to a study of Twitter data, which showed an increased in tweets expressing mental health concerns due to infection risk and isolation strategies in the early stages of the pandemic [30]. After examining 100 samples of related portal messages, we found that patients often reported worry about worsening of their current illnesses without their typical in-person follow-up as well as concern for COVID-19 infection risk during their visits to the clinics or hospitals. Some patients were also nervous about falling ill because they were not able to afford health care services owing to loss of jobs and health insurances [31]. Under such stressors, some patients reported depressive symptoms, and a few indicated suicidal thoughts and requested medical advice from their providers [32,33]. Although the number of PGMs on these issues rose substantially at the beginning of March, the relevant PGM count was progressively declining over time after April. The trends of PGMs related to the mental health concerns were consistent with those of PGMs related to COVID-19 care plans. We speculate the findings on the decline in overall mental health issues due to COVID-19 among patient portal users may result from a reduction in COVID-19 risk, eventual management of medical conditions, adaption to mitigation activities, or support from their caregivers. For example, a study on web-based search behavior for mental health concerns in the United States demonstrated a significant flattening of the curve for searches for anxiety and suicidal ideation after implementation of stay-at-home orders in certain states [34].

After analyzing PGMs related to COVID-19 and unique patient senders from the Epic Clarity system (see Figure S2 in [Multimedia Appendix 1](#)), we found a low messaging rate per patient and a strong correlation between message count and unique patient count, which suggests that the volume of unique patients mainly contribute to the intensive utilization of portal messaging. Demographic analysis of patient populations shows a significant difference in the distribution of patient populations between general message senders and general portal users. Compared to general portal users, the frequent message senders were more so middle-aged adults (30-64 years old), female, married (or with life partners), non-Hispanic or Latino, White, English speakers, or urban residents. This proportion was even more pronounced in COVID-19 message senders. Middle-aged (or female, married, White, and urban) patients were more inclined to use patient portals for addressing their issues regarding COVID-19. This phenomenon is interesting given the fact that the COVID-19 pandemic disproportionately affected racial minorities and rural populations, who are particularly vulnerable to severe outcomes of COVID-19 [35,36]. This may be a consequence of the patient population of the institution or rather that patient portal has some inherent bias toward more

health-literate patient populations, which may not be the same populations as negatively affected by the pandemic.

Telehealth, including the use of patient portals, is transforming the delivery of health care [2]. An unintended effect of the growth of patient portal messages may be an increased workload for providers. A previous study reported that providers occasionally needed to reply to messages sent by patients after work hours in order to ensure timely response. Newer delivery models will need to properly distribute the communication load for better efficiency and avoid provider burnout [26,37]. Some health care systems will probably face this challenge of managing increasing volumes of patient messages in the near future [38]. These health care systems will require new billing models and practice metrics, or additional ancillary infrastructure, including support staff, to accommodate this growing trend of asynchronous communication. Evolving technologies in artificial intelligence and natural language processing tools may even be considered as technological support for care teams in secure messaging [39,40]. In addition, the observed disparities in use of remote patient care among these populations warrants attention from providers and researchers on designing inclusive as well as innovative solutions to achieve equity in health care service delivery [41].

Limitations

There are several limitations to our study. First, the patient portal messages were collected at Mayo Clinic, a multispecialty academic medical center. The collected data might not be representative of different clinical settings or patient populations in other areas of the country. Second, keyword searching was carried out to identify patient portal messages associated with COVID-19, COVID-19-related care, and other health care issues due to COVID-19. Although the keyword sets cover a large number of relevant keywords, synonyms, and morphological variations, they may not be totally comprehensive; hence, bias could exist in our results. We are developing robust detection algorithms based on state-of-art deep learning techniques to accurately identify interesting health topics in portal messages during the pandemic. Finally, we investigated patient portal messages in the early stage of the COVID-19 pandemic. The analysis of patient portal messages in the following time or later stages of the pandemic is beyond the scope of this study but represents an important area for further exploration.

Conclusions

During the COVID-19 pandemic, patient portal utilization increased to address questions and concerns about the COVID-19 pandemic, revolving mainly around symptom self-assessment, tests, and results. The increased usage statistics for COVID-19 indicates the patient portal was a valuable web-based platform for patients to remotely discuss COVID-19 diagnosis and treatment as well as seek support for other health care issues impacted by the pandemic. The volume of PGMs on COVID-19-related care fluctuated as the pandemic developed. After initial increase in March, the PGMs regarding other health care issues such as appointment cancellations and anxiety about disease progression exhibited a declining trend. We observed differences in patient demographics between

general portal users, general message senders, and COVID-19-specific message senders, mainly that the majority demographic took on a larger proportion of COVID-19 messages. There is still great potential to increase PGM engagement for minority populations and rural communities

with regard to the COVID-19 pandemic. Time-series analysis of portal messages could offer us a timely surveillance of COVID-19 and its impacts on patients to improve patient-centered care related to the COVID-19 crisis.

Acknowledgments

This research work was supported by the National Center for Advancing Translational Science of the National Institutes of Health (award U01TR002062). The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health.

Conflicts of Interest

None declared.

Multimedia Appendix 1

Table S1 Keywords used to filter relevant portal messages; Table S2. Performance of Keyword searching methods; Table S3 Chi-squared goodness-of-fit test between patient role groups; Figure S1 Newly confirmed COVID-19 cases, hospitalized cases, and deaths in the US and Minnesota; Figure S2 Daily numbers and weekly smoothing averages (WSAs) of patient-generated messages (PGMs), unique patient senders, and messages per patient from the Epic Clarity database.

[[DOCX File, 269 KB - humanfactors_v9i2e35187_app1.docx](#)]

References

1. Budd J, Miller BS, Manning EM, Lampos V, Zhuang M, Edelstein M, et al. Digital technologies in the public-health response to COVID-19. *Nat Med* 2020 Aug;26(8):1183-1192. [doi: [10.1038/s41591-020-1011-4](https://doi.org/10.1038/s41591-020-1011-4)] [Medline: [32770165](https://pubmed.ncbi.nlm.nih.gov/32770165/)]
2. Wosik J, Fudim M, Cameron B, Gellad Z, Cho A, Phinney D, et al. Telehealth transformation: COVID-19 and the rise of virtual care. *J Am Med Inform Assoc* 2020 Jun 01;27(6):957-962 [FREE Full text] [doi: [10.1093/jamia/ocaa067](https://doi.org/10.1093/jamia/ocaa067)] [Medline: [32311034](https://pubmed.ncbi.nlm.nih.gov/32311034/)]
3. Ferguson J, Jacobs J, Yefimova M, Greene L, Heyworth L, Zulman D. Virtual care expansion in the Veterans Health Administration during the COVID-19 pandemic: clinical services and patient characteristics associated with utilization. *J Am Med Inform Assoc* 2021 Mar 01;28(3):453-462 [FREE Full text] [doi: [10.1093/jamia/ocaa284](https://doi.org/10.1093/jamia/ocaa284)] [Medline: [33125032](https://pubmed.ncbi.nlm.nih.gov/33125032/)]
4. Irizarry T, DeVito Dabbs A, Curran CR. Patient Portals and Patient Engagement: A State of the Science Review. *J Med Internet Res* 2015 Jun 23;17(6):e148 [FREE Full text] [doi: [10.2196/jmir.4255](https://doi.org/10.2196/jmir.4255)] [Medline: [26104044](https://pubmed.ncbi.nlm.nih.gov/26104044/)]
5. Kruse CS, Bolton K, Freriks G. The effect of patient portals on quality outcomes and its implications to meaningful use: a systematic review. *J Med Internet Res* 2015 Feb 10;17(2):e44 [FREE Full text] [doi: [10.2196/jmir.3171](https://doi.org/10.2196/jmir.3171)] [Medline: [25669240](https://pubmed.ncbi.nlm.nih.gov/25669240/)]
6. Henry J, Barker W, Kachay L. Electronic Capabilities for Patient Engagement among U.S. Non-Federal Acute Care Hospitals: 2013-2017. The Office of the National Coordinator for Health Information Technology. 2019. URL: <https://www.healthit.gov/sites/default/files/page/2019-04/AHApatientengagement.pdf> [accessed 2022-04-21]
7. Tieu L, Sarkar U, Schillinger D, Ralston JD, Ratanawongsa N, Pasick R, et al. Barriers and Facilitators to Online Portal Use Among Patients and Caregivers in a Safety Net Health Care System: A Qualitative Study. *J Med Internet Res* 2015 Dec 03;17(12):e275 [FREE Full text] [doi: [10.2196/jmir.4847](https://doi.org/10.2196/jmir.4847)] [Medline: [26681155](https://pubmed.ncbi.nlm.nih.gov/26681155/)]
8. Graetz I, Huang J, Muelly ER, Fireman B, Hsu J, Reed ME. Association of Mobile Patient Portal Access With Diabetes Medication Adherence and Glycemic Levels Among Adults With Diabetes. *JAMA Netw Open* 2020 Feb 05;3(2):e1921429 [FREE Full text] [doi: [10.1001/jamanetworkopen.2019.21429](https://doi.org/10.1001/jamanetworkopen.2019.21429)] [Medline: [32074289](https://pubmed.ncbi.nlm.nih.gov/32074289/)]
9. Steitz BD, Unertl KM, Levy MA. Characterizing communication patterns among members of the clinical care team to deliver breast cancer treatment. *J Am Med Inform Assoc* 2020 Feb 01;27(2):236-243 [FREE Full text] [doi: [10.1093/jamia/ocz151](https://doi.org/10.1093/jamia/ocz151)] [Medline: [31682267](https://pubmed.ncbi.nlm.nih.gov/31682267/)]
10. Pecina J, North F, Williams MD, Angstman KB. Use of an on-line patient portal in a depression collaborative care management program. *J Affect Disord* 2017 Jan 15;208:1-5. [doi: [10.1016/j.jad.2016.08.034](https://doi.org/10.1016/j.jad.2016.08.034)] [Medline: [27736688](https://pubmed.ncbi.nlm.nih.gov/27736688/)]
11. United States COVID-19 Cases, Deaths, and Laboratory Testing (NAATs) by State, Territory, and Jurisdiction. Centers for Disease Control and Prevention. URL: <https://www.cdc.gov/coronavirus/2019-ncov/cases-updates/cases-in-us.html> [accessed 2022-04-21]
12. Al-Quteimat O, Amer A. The Impact of the COVID-19 Pandemic on Cancer Patients. *Am J Clin Oncol* 2020 Jun;43(6):452-455. [doi: [10.1097/COC.0000000000000712](https://doi.org/10.1097/COC.0000000000000712)] [Medline: [32304435](https://pubmed.ncbi.nlm.nih.gov/32304435/)]
13. Mehrotra A, Chernew M, Linetsky D, Hatch H, Cutler D. The Impact of the COVID-19 Pandemic on Outpatient Visits: A Rebound Emerges. The Commonwealth Fund. 2020 May 19. URL: <https://www.commonwealthfund.org/publications/2020/apr/impact-covid-19-outpatient-visits> [accessed 2022-04-21]

14. Hartnett KP, Kite-Powell A, DeVies J, Coletta MA, Boehmer TK, Adjemian J, National Syndromic Surveillance Program Community of Practice. Impact of the COVID-19 Pandemic on Emergency Department Visits - United States, January 1, 2019-May 30, 2020. *MMWR Morb Mortal Wkly Rep* 2020 Jun 12;69(23):699-704 [[FREE Full text](#)] [doi: [10.15585/mmwr.mm6923e1](https://doi.org/10.15585/mmwr.mm6923e1)] [Medline: [32525856](https://pubmed.ncbi.nlm.nih.gov/32525856/)]
15. Mann D, Chen J, Chunara R, Testa P, Nov O. COVID-19 transforms health care through telemedicine: Evidence from the field. *J Am Med Inform Assoc* 2020 Jul 01;27(7):1132-1135 [[FREE Full text](#)] [doi: [10.1093/jamia/ocaa072](https://doi.org/10.1093/jamia/ocaa072)] [Medline: [32324855](https://pubmed.ncbi.nlm.nih.gov/32324855/)]
16. Rajkumar RP. COVID-19 and mental health: A review of the existing literature. *Asian J Psychiatr* 2020 Aug;52:102066 [[FREE Full text](#)] [doi: [10.1016/j.ajp.2020.102066](https://doi.org/10.1016/j.ajp.2020.102066)] [Medline: [32302935](https://pubmed.ncbi.nlm.nih.gov/32302935/)]
17. Lai J, Ma S, Wang Y, Cai Z, Hu J, Wei N, et al. Factors Associated With Mental Health Outcomes Among Health Care Workers Exposed to Coronavirus Disease 2019. *JAMA Netw Open* 2020 Mar 02;3(3):e203976 [[FREE Full text](#)] [doi: [10.1001/jamanetworkopen.2020.3976](https://doi.org/10.1001/jamanetworkopen.2020.3976)] [Medline: [32202646](https://pubmed.ncbi.nlm.nih.gov/32202646/)]
18. Torales J, O'Higgins M, Castaldelli-Maia JM, Ventriglio A. The outbreak of COVID-19 coronavirus and its impact on global mental health. *Int J Soc Psychiatry* 2020 Jun;66(4):317-320. [doi: [10.1177/0020764020915212](https://doi.org/10.1177/0020764020915212)] [Medline: [32233719](https://pubmed.ncbi.nlm.nih.gov/32233719/)]
19. Pfefferbaum B, North CS. Mental Health and the Covid-19 Pandemic. *N Engl J Med* 2020 Aug 06;383(6):510-512. [doi: [10.1056/nejmp2008017](https://doi.org/10.1056/nejmp2008017)]
20. Panchal N, Kamal R, Orgera K, Cox C, Garfield R, Hamel L. The Implications of COVID-19 for Mental Health and Substance Use. Kaiser Family Foundation. 2021 Feb 10. URL: <https://www.kff.org/coronavirus-covid-19/issue-brief/the-implications-of-covid-19-for-mental-health-and-substance-use/> [accessed 2022-04-21]
21. Twenge JM, Joiner TE. Mental distress among U.S. adults during the COVID-19 pandemic. *J Clin Psychol* 2020 Dec;76(12):2170-2182 [[FREE Full text](#)] [doi: [10.1002/jclp.23064](https://doi.org/10.1002/jclp.23064)] [Medline: [33037608](https://pubmed.ncbi.nlm.nih.gov/33037608/)]
22. Lee I, Kovarik C, Tejasvi T, Pizarro M, Lipoff JB. Telehealth: Helping your patients and practice survive and thrive during the COVID-19 crisis with rapid quality implementation. *J Am Acad Dermatol* 2020 May;82(5):1213-1214 [[FREE Full text](#)] [doi: [10.1016/j.jaad.2020.03.052](https://doi.org/10.1016/j.jaad.2020.03.052)] [Medline: [32229278](https://pubmed.ncbi.nlm.nih.gov/32229278/)]
23. Patel PD, Cobb J, Wright D, Turer R, Jordan T, Humphrey A, et al. Rapid development of telehealth capabilities within pediatric patient portal infrastructure for COVID-19 care: barriers, solutions, results. *J Am Med Inform Assoc* 2020 Jul 01;27(7):1116-1120 [[FREE Full text](#)] [doi: [10.1093/jamia/ocaa065](https://doi.org/10.1093/jamia/ocaa065)] [Medline: [32302395](https://pubmed.ncbi.nlm.nih.gov/32302395/)]
24. Khairat S, McDaniel P, Jansen M, Francis T, Edson B, Gianforcaro R. Analysis of Social Determinants and the Utilization of Pediatric Tele-Urgent Care During the COVID-19 Pandemic: Cross-sectional Study. *JMIR Pediatr Parent* 2021 Aug 30;4(3):e25873 [[FREE Full text](#)] [doi: [10.2196/25873](https://doi.org/10.2196/25873)] [Medline: [34459742](https://pubmed.ncbi.nlm.nih.gov/34459742/)]
25. Portz JD, Brungardt A, Shanbhag P, Staton EW, Bose-Brill S, Lin C, et al. Advance Care Planning Among Users of a Patient Portal During the COVID-19 Pandemic: Retrospective Observational Study. *J Med Internet Res* 2020 Aug 11;22(8):e21385 [[FREE Full text](#)] [doi: [10.2196/21385](https://doi.org/10.2196/21385)] [Medline: [32716900](https://pubmed.ncbi.nlm.nih.gov/32716900/)]
26. Nath B, Williams B, Jeffery MM, O'Connell R, Goldstein R, Sinsky CA, et al. Trends in Electronic Health Record Inbox Messaging During the COVID-19 Pandemic in an Ambulatory Practice Network in New England. *JAMA Netw Open* 2021 Oct 01;4(10):e2131490 [[FREE Full text](#)] [doi: [10.1001/jamanetworkopen.2021.31490](https://doi.org/10.1001/jamanetworkopen.2021.31490)] [Medline: [34636917](https://pubmed.ncbi.nlm.nih.gov/34636917/)]
27. Ioannidis JP, Axfors C, Contopoulos-Ioannidis DG. Second versus first wave of COVID-19 deaths: Shifts in age distribution and in nursing home fatalities. *Environ Res* 2021 Apr;195:110856 [[FREE Full text](#)] [doi: [10.1016/j.envres.2021.110856](https://doi.org/10.1016/j.envres.2021.110856)] [Medline: [33581086](https://pubmed.ncbi.nlm.nih.gov/33581086/)]
28. Patient Online Services. Mayo Clinic. URL: <https://onlineservices.mayoclinic.org/content/staticpatient/showpage/patientonline> [accessed 2022-04-21]
29. Denis F, Fontanet A, Le Douarin Y, Le Goff F, Jeanneau S, Lescure F. A Self-Assessment Web-Based App to Assess Trends of the COVID-19 Pandemic in France: Observational Study. *J Med Internet Res* 2021 Mar 12;23(3):e26182 [[FREE Full text](#)] [doi: [10.2196/26182](https://doi.org/10.2196/26182)] [Medline: [33709945](https://pubmed.ncbi.nlm.nih.gov/33709945/)]
30. Valdez D, Ten Thij M, Bathina K, Rutter LA, Bollen J. Social Media Insights Into US Mental Health During the COVID-19 Pandemic: Longitudinal Analysis of Twitter Data. *J Med Internet Res* 2020 Dec 14;22(12):e21418 [[FREE Full text](#)] [doi: [10.2196/21418](https://doi.org/10.2196/21418)] [Medline: [33284783](https://pubmed.ncbi.nlm.nih.gov/33284783/)]
31. Woolhandler S, Himmelstein DU. Intersecting U.S. Epidemics: COVID-19 and Lack of Health Insurance. *Ann Intern Med* 2020 Jul 07;173(1):63-64 [[FREE Full text](#)] [doi: [10.7326/M20-1491](https://doi.org/10.7326/M20-1491)] [Medline: [32259195](https://pubmed.ncbi.nlm.nih.gov/32259195/)]
32. Gunnell D, Appleby L, Arensman E, Hawton K, John A, Kapur N, COVID-19 Suicide Prevention Research Collaboration. Suicide risk and prevention during the COVID-19 pandemic. *Lancet Psychiatry* 2020 Jun;7(6):468-471 [[FREE Full text](#)] [doi: [10.1016/S2215-0366\(20\)30171-1](https://doi.org/10.1016/S2215-0366(20)30171-1)] [Medline: [32330430](https://pubmed.ncbi.nlm.nih.gov/32330430/)]
33. Reger MA, Stanley IH, Joiner TE. Suicide Mortality and Coronavirus Disease 2019-A Perfect Storm? *JAMA Psychiatry* 2020 Nov 01;77(11):1093-1094. [doi: [10.1001/jamapsychiatry.2020.1060](https://doi.org/10.1001/jamapsychiatry.2020.1060)] [Medline: [32275300](https://pubmed.ncbi.nlm.nih.gov/32275300/)]
34. Jacobson NC, Lekkas D, Price G, Heinz MV, Song M, O'Malley AJ, et al. Flattening the Mental Health Curve: COVID-19 Stay-at-Home Orders Are Associated With Alterations in Mental Health Search Behavior in the United States. *JMIR Ment Health* 2020 Jun 01;7(6):e19347 [[FREE Full text](#)] [doi: [10.2196/19347](https://doi.org/10.2196/19347)] [Medline: [32459186](https://pubmed.ncbi.nlm.nih.gov/32459186/)]
35. Kirby T. Evidence mounts on the disproportionate effect of COVID-19 on ethnic minorities. *Lancet Respir Med* 2020 Jun;8(6):547-548 [[FREE Full text](#)] [doi: [10.1016/S2213-2600\(20\)30228-9](https://doi.org/10.1016/S2213-2600(20)30228-9)] [Medline: [32401711](https://pubmed.ncbi.nlm.nih.gov/32401711/)]

36. Peters D. Rural Areas Face Higher and Distinct Risks of Serious COVID-19 Outcomes than Urban Areas. *Small* 2020;999k.
37. Huang M, Fan J, Prigge J, Shah N, Costello B, Yao L. Characterizing Patient-Clinician Communication in Secure Medical Messages: Retrospective Study. *J Med Internet Res* 2022 Jan 11;24(1):e17273 [FREE Full text] [doi: [10.2196/17273](https://doi.org/10.2196/17273)] [Medline: [35014964](https://pubmed.ncbi.nlm.nih.gov/35014964/)]
38. Hoonakker PL, Carayon P, Cartmill RS. The impact of secure messaging on workflow in primary care: Results of a multiple-case, multiple-method study. *Int J Med Inform* 2017 Apr;100:63-76 [FREE Full text] [doi: [10.1016/j.ijmedinf.2017.01.004](https://doi.org/10.1016/j.ijmedinf.2017.01.004)] [Medline: [28241939](https://pubmed.ncbi.nlm.nih.gov/28241939/)]
39. Sulieman L, Gilmore D, French C, Cronin RM, Jackson GP, Russell M, et al. Classifying patient portal messages using Convolutional Neural Networks. *J Biomed Inform* 2017 Oct;74:59-70 [FREE Full text] [doi: [10.1016/j.jbi.2017.08.014](https://doi.org/10.1016/j.jbi.2017.08.014)] [Medline: [28864104](https://pubmed.ncbi.nlm.nih.gov/28864104/)]
40. Sulieman L, Robinson JR, Jackson GP. Automating the Classification of Complexity of Medical Decision-Making in Patient-Provider Messaging in a Patient Portal. *J Surg Res* 2020 Nov;255:224-232 [FREE Full text] [doi: [10.1016/j.jss.2020.05.039](https://doi.org/10.1016/j.jss.2020.05.039)] [Medline: [32570124](https://pubmed.ncbi.nlm.nih.gov/32570124/)]
41. Crawford A, Serhal E. Digital Health Equity and COVID-19: The Innovation Curve Cannot Reinforce the Social Gradient of Health. *J Med Internet Res* 2020 Jun 02;22(6):e19361 [FREE Full text] [doi: [10.2196/19361](https://doi.org/10.2196/19361)] [Medline: [32452816](https://pubmed.ncbi.nlm.nih.gov/32452816/)]

Abbreviations

MCHS: Mayo Clinic Health System

PGM: patient-generated message

WSA: weekly smoothing average

Edited by A Kushniruk; submitted 24.11.21; peer-reviewed by Q Song, H Ayatollahi, R Menhas; comments to author 07.12.21; revised version received 09.01.22; accepted 14.02.22; published 05.05.22.

Please cite as:

Huang M, Khurana A, Mastorakos G, Wen A, He H, Wang L, Liu S, Wang Y, Zong N, Prigge J, Costello B, Shah N, Ting H, Fan J, Patten C, Liu H

Patient Portal Messaging for Asynchronous Virtual Care During the COVID-19 Pandemic: Retrospective Analysis

JMIR Hum Factors 2022;9(2):e35187

URL: <https://humanfactors.jmir.org/2022/2/e35187>

doi: [10.2196/35187](https://doi.org/10.2196/35187)

PMID: [35171108](https://pubmed.ncbi.nlm.nih.gov/35171108/)

©Ming Huang, Aditya Khurana, George Mastorakos, Andrew Wen, Huan He, Liwei Wang, Sijia Liu, Yanshan Wang, Nansu Zong, Julie Prigge, Brian Costello, Nilay Shah, Henry Ting, Jungwei Fan, Christi Patten, Hongfang Liu. Originally published in *JMIR Human Factors* (<https://humanfactors.jmir.org>), 05.05.2022. This is an open-access article distributed under the terms of the Creative Commons Attribution License (<https://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in *JMIR Human Factors*, is properly cited. The complete bibliographic information, a link to the original publication on <https://humanfactors.jmir.org>, as well as this copyright and license information must be included.

Original Paper

Modeling Trust in COVID-19 Contact-Tracing Apps Using the Human-Computer Trust Scale: Online Survey Study

Sonia Sousa^{1,2*}, PhD; Tiina Kalju^{1*}, MSc

¹School of Digital Technologies, Tallinn University, Tallinn, Estonia

²Institute for Systems and Computer Engineering, Technology and Science, University of Trás-os-Montes and Alto Douro, Vila Real, Portugal

* all authors contributed equally

Corresponding Author:

Sonia Sousa, PhD

School of Digital Technologies

Tallinn University

Narva mnt, 29

Tallinn, 10120

Estonia

Phone: 372 53921116

Email: scs@tlu.ee

Abstract

Background: The COVID-19 pandemic has caused changes in technology use worldwide, both socially and economically. This pandemic crisis has brought additional measures such as contact-tracing apps (CTAs) to help fight against spread of the virus. Unfortunately, the low adoption rate of these apps affected their success. There could be many reasons for the low adoption, including concerns of security and privacy, along with reported issues of trust in CTAs. Some concerns are related with how CTAs could be used as surveillance tools or their potential threats to privacy as they involve health data. For example, in Estonia, the CTA named HOIA had approximately 250,000 downloads in the middle of January 2021. However, in 2021, only 4.7% of the population used HOIA as a COVID-19 CTA. The reasons for the low adoption include lack of competency, and privacy and security concerns. This lower adoption and the lack of trustworthiness persist despite efforts of the European Union in building ethics and trustworthy artificial intelligence (AI)-based apps.

Objective: The aim of this study was to understand how to measure trust in health technologies. Specifically, we assessed the usefulness of the Human-Computer Trust Scale (HCTS) to measure Estonians' trust in the HOIA app and the causes for this lack of trust.

Methods: The main research question was: Can the HCTS be used to assess citizens' perception of trust in health technologies? We established four hypotheses that were tested with a survey. We used a convenience sample for data collection, including sharing the questionnaire on social network sites and using the snowball method to reach all potential HOIA users in the Estonian population.

Results: Among the 78 respondents, 61 had downloaded the HOIA app with data on usage patterns. However, 20 of those who downloaded the app admitted that it was never opened despite most claiming to regularly use mobile apps. The main reasons included not understanding how it works, and privacy and security concerns. Significant correlations were found between participants' trust in CTAs in general and their perceived trust in the HOIA app regarding three attributes: competency ($P < .001$), risk perception ($P < .001$), and reciprocity ($P = .01$).

Conclusions: This study shows that trust in the HOIA app among Estonian residents did affect their predisposition to use the app. Participants did not generally believe that HOIA could help to control the spread of the virus. The result of this work is limited to HOIA and health apps that use similar contact-tracing methods. However, the findings can contribute to gaining a broader understanding and awareness of the need for designing trustworthy technologies. Moreover, this work can help to provide design recommendations that ensure trustworthiness in CTAs, and the ability of AI to use highly sensitive data and serve society.

(JMIR Hum Factors 2022;9(2):e33951) doi:[10.2196/33951](https://doi.org/10.2196/33951)

KEYWORDS

human-computer interaction; COVID-19; human factors; trustworthy AI; contact-tracing; app; safety; trust; artificial intelligence; Estonia; case study; monitoring; surveillance; perspective; awareness; design; covid; mobile app; mHealth; mobile health

Introduction

Background

The COVID-19 pandemic has changed how we view technology as a resource to stop the spread of disease. To address the need to control the spread of the virus, many governments and public health authorities worldwide have launched several technological initiatives, including the development of artificial intelligence (AI) contact-tracing mobile apps (CTAs). As a result, by the end of 2020, there were more than 50 CTAs available in both Google Play and iOS App Store [1,2]. According to Nguyen et al [3], security and privacy are crucial in designing AI-based CTA technologies. If users perceive CTAs as a threat to their privacy, this might affect their predisposition to use the app, ultimately affecting its adoption rate and tool effectiveness. This evidence has led to an increased discourse for design systems toward focusing on ensuring that CTAs are secure and private. Previous studies have recommended several criteria such as ensuring a low level of complexity of the security feature so that it is easy to use and understandable for the general population [4,5], visibility and interaction from the user, and unambiguous and clear messages to follow while designing security measures [6-8]. Similar arguments were put forth in Europe's stated goals to ensure ethical and responsible technological development. Although COVID-19 CTAs in Europe followed the General Data Protection Regulation and ISO/IEC 27001 [9] regulations, and were also designed in consideration of current AI principles to regulate technology use (ie, Ethical guidelines for Trustworthy AI [10]), this was not sufficient to ensure the trustworthiness from citizens. This lack of trustworthiness exists despite widely available information on how these technologies were built with transparent and ethical principles in mind. Moreover, despite government initiatives to push through their adoption, the download rates and actual usage rates of these apps remained low [2,6,11-13]. One reason for this low adoption might be that security and privacy in computer science are still mainly approached from a technical perspective [14]. Privacy attributes in technology can be more profound and complex than technical qualities. Privacy is defined as a person's control over the information that is manipulated and communicated to others [6,15-18].

Privacy also includes interpersonal characteristics such as the perception of privacy, system honesty or benevolence communication, and shared control to minimize associated risk and uncertainty. For instance, despite appropriate regulations and principles being considered when designing Estonia's COVID-19 CTA (HOIA), the adoption of HOIA by citizens did not increase. The critical reasons for the low adoption of HOIA included lack of effectiveness (10%) and concerns of security and privacy (19%) according to a survey initiated by The Ministry of Social Affairs, surveying 92% of Estonian residents [13,19]. Thus, all efforts made in designing AI-based transparent and ethically responsible CTAs that can prevent

data misuse and ensure the development of responsible trustworthy AI interactions were unsuccessful.

We believe that it is essential to find new ways to ensure incorporating trust values in the design of such apps that could lead to building more technological, socially responsible societies. One should expect trust to be increasingly in demand as a means of enduring the complexity of a future that technology will generate. The quality and depth of technology use are also significantly affected by users' trust in the technology. Trust is defined according to the ability to determine who to trust, and represents the willingness of a party to be vulnerable to the actions of another party based on the expectation that the other will perform a particular action important to the trustor, irrespective of the ability to monitor or control that other party [20-22].

Research Gaps

Prior research confirms that technology acceptance and adoption are affected by the level of trust users have in the technology [11,20-23]. However, evidence shows that designing trustworthy technologies is complex and needs to be better understood. Like privacy, trust is an interpersonal quality that is present in many moments of our daily lives, and is thus often considered unconsciously. Whether being conscious or unconscious of its existence, trust represents an important key of the relationships encountered in daily life, including interactions between humans and machines. Establishing a trustful relationship implies peoples' permission to share knowledge, delegation, and cooperative actions [11,22,24,25]. Thus, in addition to the current research challenge for ensuring that all ethical, privacy, and technical security requirements are considered [5,7,9], we argue that trust might be the reason why users do not feel comfortable using CTAs that depend on citizens' data to function properly. If this is indeed the case, besides existing design regulations and principles, designers will also need mechanisms to analyze individuals' perceived trustworthiness in AI apps. In this way, designers and other stakeholders can gain a deeper understanding of how individuals perceive the benefits of AI, and assess their predisposition to cooperate and be more willing to use the technologies. Thus, it is important to gauge the extent to which such AI data-driven technologies are perceived as trustworthy (ie, the gains of using CTAs are higher than the possible losses).

There are three main rationales for the above argument. First, with the current culture of increased introduction and use of complex systems in our daily activities, researchers need to focus more on conceiving responsible human-computer interactions. Second, current paradigms supporting ethical and responsible design practices are insufficient to ensure technology trustworthiness. Third, a new human-machine interaction mechanism is needed to effectively evaluate users' trust perceptions in technology (eg, assess users' experience toward incorporated trust values). Namely, we propose new human-centered design frameworks and mechanisms to guide

the design and technology evaluation process. Overall, in the past decade, human-computer interaction has contributed significantly toward improving the quality of living with technology. Consequently, regular individuals are getting more involved, engaged, and dependent on technology to achieve their goals. It is true that we no longer live without technology. Despite this, the above arguments indicate that we are entering a new era that depends on data to thrive. This symbiotic dependence of humans in systems abilities and of systems dependence in our data to provide meaningful information has increased the complexity of the technology provided. Consequently, we have become more reliant on trust to survive in these complex symbiotic relationships. This is clearly shown in how digital CTAs were affected by these symbiotic relationships. Most of these apps are collecting highly sensitive data from individuals, including where they have been and with whom they have been in contact.

Methods

Study Aims and Design

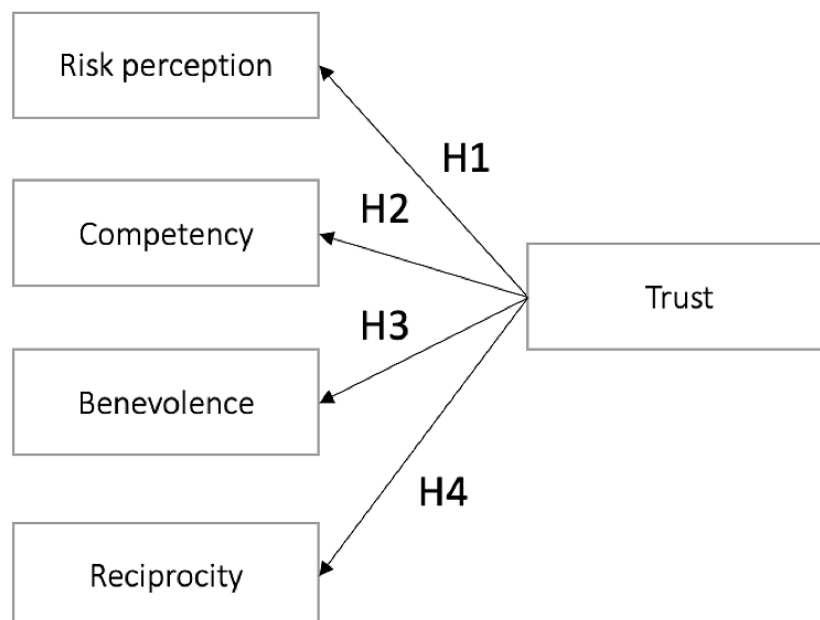
This study builds on the prior work of Gulati et al [20] and Sousa et al [22], and is guided by one central research question: Can the Human-Computer Trust Scale (HCTS) be used to assess an individual's perception of trust in health technologies? The main goal of this study was to propose a novel design evaluation mechanism to incorporate trust values in health care technologies, and make health care interventions and technologies more trustworthy and accepted. Namely, we used partial least-squares structural equation modeling (PLS-SEM)

to empirically ascertain which attributes of the proposed scale (HCTS) hold in health care contexts and can be used as lenses to evaluate individuals' trust predisposition to interact. The study was divided into two main stages: (1) adaptation and translation of the scale, and (2) measurement and validation of the questionnaire (HCTS).

Theoretical Model

The adopted theoretical model, the HCTS [20], illustrates the multidimensional nature of trust, taking into account several attributes of trust, as shown in [Figure 1](#). This model was validated with statistical modeling techniques. The proposed attributes of the model were gathered from a systematic multidisciplinary literature review, combined with (1) a word elicitation study to capture a rich set of multidisciplinary notions encapsulating trust; (2) participatory design sessions and exploratory interviews with users to further identify antecedents of trust; (3) the unification of technology acceptance models [22]; and (4) separate studies to ensure statistical certainty of the scale proposed: trust in Siri, trust in the Estonian electronic voting system, trust in futuristic scenarios, and trust in human-robot interaction [20,26]. The final scale to measure trust consists of three main attributes: risk perception, competency, and benevolence. In line with the above findings and with the awareness that trust assessment is context- and culture-dependent, we assessed the validity of the scale to measure citizens' trust attitudes in CTAs. To achieve our goal, we developed four sets of assumptions that might affect or predict a user's trust when interacting with the HOIA app. The four hypotheses (H1-H4) established in regard to our main research question are outlined in [Textbox 1](#).

Figure 1. Human-computer trust model under investigation. H: Hypothesis.



Textbox 1. Hypotheses of the study.**Hypothesis 1**

- There is a significant and positive association between risk perception in the HOIA app and general trust in HOIA. Risk perception is defined as *the extent to which one party is willing to participate in a given action while considering the risk and incentives involved*. Here, we assumed that the extent to which individuals are willing to participate in a given action (ie, to use HOIA) while considering that the risk and incentives involved are directly associated with their perception of technology trustworthiness: with a higher perceived risk, there will be less willingness to interact; with a lower perceived risk, users will be more willing to interact.

Hypothesis 2

- There is a significant and positive association between competence and general trust in HOIA. HOIA competence is defined as *the ease of use associated with the use of a system in that it is perceived to perform its tasks accurately and correctly*. Here, we assumed that an individual's perception of a contact-tracing app as competent is based on its functionality, closely linked to the concept of usefulness of a system. Higher perceived competency indicates that participants perceived the tool to be capable of doing what is expected, be useful, and will help them achieve desired goals.

Hypothesis 3

- There is a significant and positive association between benevolence and general trust in HOIA. Benevolence is defined as a citizen's *perception that a particular system will act in their best interest and that most people using the system share similar social behaviors and values*. Here, we assumed that an individual's perception that a particular system will act in their best interest, and that most people using the system share similar social behaviors and values that a particular technology will provide. Higher perceptions of benevolence are associated with fewer risks and uncertainties in its use.

Hypothesis 4

- There is a significant and positive association between reciprocity and trust in HOIA use. The notion of reciprocity is understood as *the degree to which an individual sees oneself as a part of a group*. It is built on the principle of mutual benefit, feeling a sense of belonging, and feeling connected, based on the give-and-take principles associated with the notion of computers as social actors. Here, we assumed that a citizen's perception of contact tracing apps is reciprocal based on the degree to which an individual sees oneself as a part of a group.

Study Procedure**Questionnaire**

We used a semistructured questionnaire to collect data. Before distributing the questionnaire, we adapted the original scale to the context and translated the content from English into Estonian. The translation and adaptation of the instrument followed the guidelines of the adaptation, translation, and validation process [27]. The survey was designed based on the HCTS in the Estonian language and was administered during April 2021. The objective of this study was to build on prior works and empirically assess HCTS to ascertain which attributes of the model hold true in health user–technology interactions.

The survey was created using both Lime Survey and Google Forms. During the pilot study, the feedback from the respondents was that the visual design of the Google Forms is less confusing; therefore, it was decided to adopt Google Forms as the final survey format.

Stimuli

To ensure that all participants understood the technical artefact in question and their perceptions of trust regarding similar experiences, we provided the official video that explains HOIA to the users as a stimulus, following the concept of technology probe and design fiction, also known as a vignette-based study in psychology.

Recruitment

The survey was carried out among the Estonian population, which was distributed online, mainly through Facebook and other social network groups available to the authors. A

convenience sample was used in data collection because this enables reaching members of the population who are easily accessible, available, and willing to participate [28].

Ethical Considerations

This study complies with the basic ethical principles for the responsible conduct of research involving human subjects. Informed consent was requested from all participants, and authorization was obtained from the authors of the scale [20] to carry out the contextual adaptation and validation of the scale. The study was approved by the Tallinn University Ethics committee on July 9th, 2021 (study name: "Survey on the dynamic trust relationships between technology, society and culture"; approval number: Taotlus nr 6-5.1/17).

Results**Participant Characteristics**

A total of 78 responses were obtained and used for data analyses; very few responses were excluded as all respondents fully completed the survey. The three excluded cases included answers leaning in majority toward neutral options. Data collected included the following information: demographics, usage patterns of mobile apps and HOIA, trust in HOIA (including risk perception, benevolence, competence, and general trust), and opinions about HOIA's existing and additional functionalities. Among the 78 respondents, 73% (n=57) were women and only 27% (n=21) were men. Almost half of the respondents (36/78, 47%) were between the ages of 31-42 years and approximately one third (25/78, 32%) were 43-55 years old.

HOIA Usage Patterns

Among the 78 respondents, 61 had downloaded the HOIA CTA. Among them, the 47 women showed the highest rate of downloads compared with the 14 male respondents. Younger respondents (aged 18-30 years) had a higher number of downloads (88%), but they also represented the smallest sample. Slightly more than half of the participants (56%) admitted that they do not feel confident in how to use HOIA; this perception was more prominent among men (n=13). Twenty participants admitted that they had never opened the app, despite 61 claiming to use mobile apps daily.

Among the 17 respondents who had not downloaded the HOIA app, the majority were men. The main reasons claimed by participants for not downloading HOIA included: do not understand how it works, and concerns about the privacy and security of their data. When asked what additional features they expect from the CTA, some mentioned the need to understand the benefits of using it actively. When asked about their most common activities on their mobile devices, 76 participants stated that they are used for communication, 66 stated social networking, 60 stated entertainment purposes, and 40 indicated uses related to health and well-being.

Assessment of the Scale

The HCTS under investigation includes five constructs: risk perception, competency, benevolence, reciprocity, and trust

Textbox 2. Human-Computer Trust Scale measures.

Risk perception

RP1: I believe that there could be negative consequences from using HOIA

RP2: I feel I must be cautious when using HOIA

RP3: It is risky to interact with HOIA

RP4: I feel unsafe to interact with HOIA

RP5: I feel vulnerable when I interact with HOIA

Competency

COM1: I believe HOIA is competent and effective in identifying if I have been in close contact with a COVID-19-positive person

COM2: I believe HOIA has all the functionalities I would expect from a COVID-19 contact-tracing system

COM3: I believe that HOIA performs its role as a warning for close contacts with a COVID-19-positive person

Reciprocity

REC1: When I share something with HOIA, I expect to get back a knowledgeable and meaningful response

REC2: When sharing something with HOIA I believe that I will get an answer

Benevolence

BEN1: I believe HOIA acts in my best interest

BEN2: I believe that HOIA would do its best to help me if I need help

BEN3: I believe that HOIA is interested in understanding my needs and preferences

General trust

GT1: When I use HOIA, I feel I can depend on it completely

GT2: I can always rely on HOIA for guidance and assistance

GT3: I can trust the information presented to me by HOIA

[20,22,26] (see [Figure 1](#)). Following the recommendation of Hair et al [29], the minimum sample size needed to effectively perform a PLS-SEM for our study was calculated to be 40 (ie, 10 times the maximum number of arrowheads pointing at a latent variable in a PLS path model). This method was selected because measuring trust in technology is complex, including four constructs and model relationships in this case. The measures used in the study were adapted from Gulati et al [20]. Their work models trust in technology with different studies, including trust in Siri using design fiction (future scenarios), the Estonian electronic voting service, and trust in human-robot interactions [24]. Gulati et al [20] measured risk perception using the concept of willingness and motivation developed through two independent studies [6,24]. This study added two additional items created through Schoorman et al's [21] conceptualizations of trust. Gulati et al [20] measured competency and reciprocity based on the methodology of McKnight et al [30], and measured benevolence based on adaptation of the prior work of Harwood and Garry [31] and McKnight et al [30]. The survey used a 7-point Likert scale to collect data, where 1 indicates strongly disagree and 7 indicates strongly agree. All of the items were positively worded except for the risk perception scale, which was adapted as a negatively worded statement and reversed before analyzing the data. The HCTS measures are summarized in [Textbox 2](#).

Data Analysis

We analyzed a total of 78 answers. All scales for analyzing data in our study were positively worded, except perceived risk, which was negatively worded. The first steps in the analyses involved assessing the reliability and validity of the HCTS to measure trust in HOIA. In this phase, we calculated if the items have good measurements of the latent construct [29,32]. We discarded risk perception item 6 and competency item 4 because the loadings were below 0.5, and kept all loadings above their respective thresholds (>0.5). Table 1 and Figure 2 demonstrate all items used in the analysis and their loadings.

We further verified if the average variance extracted (AVE) was higher than 0.5; as shown in Table 1, all AVE values were

>0.5, demonstrating that the items have good convergent reliability [12,32]. Similarly, the composite reliability of all indicators was above >0.7, showing adequate internal consistency. The Dillon-Goldstein ρ statistic, according to Hair et al [29], is similar to Cronbach α but allows the indicator variables to have varying outer loadings, and should be higher than 0.7 (or >0.6 in exploratory research). These values were above 0.7 for all items (Table 1), further demonstrating that the model is acceptable and has satisfactory internal consistency.

The discriminant validity and cross-loading values obtained using the Fornell-Lacker criterion (Table 2) indicated that the validity of each construct is higher for itself than for each corresponding construct [32].

Table 1. Loadings, reliability, and validity of the measurement model.

Items	Loadings (>0.5)	AVE ^a (>0.5)	CR ^b (>0.7)	Dillon-Goldstein ρ (>0.7)
Benevolence		0.684	0.866	0.787
BEN1	0.780			
BEN2	0.905			
BEN3	0.791			
Competence		0.784	0.916	0.864
COM1	0.887			
COM2	0.904			
COM3	0.865			
Reciprocity		0.773	0.872	0.719
REC1	0.898			
REC2	0.860			
Risk perception		0.504	0.835	0.810
RP1	0.649			
RP2	0.727			
RP3	0.711			
RP4	0.741			
RP5	0.717			
Trust		0.622	0.830	0.717
GT1	0.822			
GT2	0.692			
GT3	0.843			

^aAVE: average variance extracted.

^bCR: composite reliability.

Figure 2. Final theoretical model loadings. BEN: benevolence; COM: competence; GT: general trust; REC: reciprocity; rev: reverse; RP: risk perception.

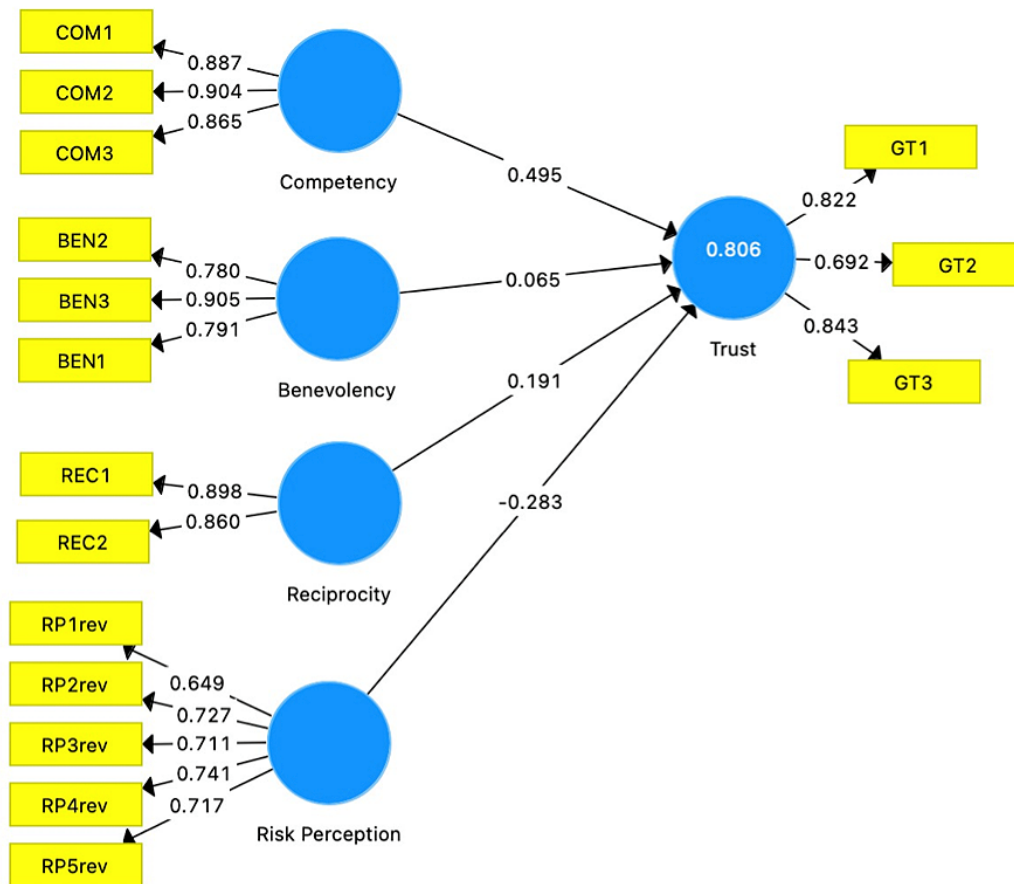


Table 2. Discriminant validity and cross-loading values (diagonal, italics) of the measurement items based on the Fornell-Lacker criterion.

Item	Benevolence	Competence	Reciprocity	Risk perception	Trust
Benevolence	<i>0.827</i>	0.747	0.620	-0.625	0.730
Competence	0.747	<i>0.885</i>	0.700	-0.585	0.843
Reciprocity	0.620	0.700	<i>0.879</i>	-0.526	0.727
Risk perception	-0.625	-0.585	-0.526	<i>0.710</i>	-0.714
Trust	0.730	0.843	0.727	-0.714	<i>0.789</i>

Trust Toward HOIA

In addition, we assessed the coefficient of determination (R^2) values, which represent the combined effect of exogenous latent variables on the endogenous latent variable, and is interpreted in the same way as in a conventional regression analysis procedure [29]. In this study, the R^2 value was 0.806 and the adjusted R^2 was 0.795. According to Hair et al [29], R^2 values of 0.75, 0.50, or 0.25 are considered substantial, moderate, or weak, respectively. In line with this interpretation, both the R^2 and adjusted R^2 values of this study indicate a substantial effect. Thus, approximately 83% of the changes in technology trust can be explained by the statistically significant exogenous variables in the HCTS. Accordingly, we conclude that the statistically significant attributes significantly predict user trust in COVID-19 CTAs, namely HOIA. Keeping in mind all of the

empirical values obtained thus far, it is safe to say that our model passes the criteria for both measurement and structural model evaluation, and the final scale exhibits good validity, reliability, and predictive power.

Discussion

Principal Findings

To contribute toward our central research question (can the HCTS be used to assess an individual’s perception of trust in health technologies?), we empirically assessed the suitability of the HCTS to assess an individual’s perception of trust in health technologies, with the broader goal of understanding which attributes of the HCTS hold true in health technologies. As shown in Table 3, all but one of our four hypotheses were supported, based on statistically significant effects.

Table 3. Significance testing of structural model path coefficients.

Hypothesis	Path coefficient (SD)	<i>t</i> value	<i>P</i> value	97.5% CI	Significance (<i>P</i> <.0.5)
Benevolence mediates trust	0.062 (0.097)	0.674	.50	0.251	No
Competency mediates trust	0.495 (0.099)	5.022	<.001	0.690	Yes
Reciprocity mediates trust	0.195 (0.084)	2.285	.02	0.355	Yes
Risk perception mediates trust	-0.287 (0.056)	5.106	<.001	-0.197	Yes

For instance, H1 (risk perception mediates trust), H2 (competency mediates trust), and H4 (reciprocity mediates trust) were statistically significant, which is in line with the work of Gulati et al [20]. However, we also found that H3 (benevolence mediates trust) was nonsignificant (*P*=.52). To understand these results, it is important to consider how these constructs were operationalized. H1 and H2 were operationalized based on Gulati et al's [20] and Schoorman et al's [21] conceptualizations of trust, whereas H3 and H4 were operationalized based on Gulati et al [20].

Limitations

Our study is not without its limitations, which can guide future research. First, culture influences trust. Second, the proposed scale (HCTS) demonstrated that trust is a dynamic construct that evolves in context and is culturally dependent. Third, the additional suggested items based on Schoorman et al's [21] conceptualizations need further reassessment, as the results are

more in line with those of Gulati et al [20], but also indicate no significant correlation between the Estonian citizens' perception of HOIA as a benevolent trait.

Conclusion

In conclusion, the results of this study indicate that the degree of trust toward the Estonian CTA (HOIA) is significantly correlated with the extent to which users perceive the system as competent, reciprocal, and risky. This study used PLS-SEM to identify statistically significant factors for assessing individuals' perception of trust in human-technology interactions for health. This work contributes toward establishing a final version of the scale derived from the HCTS consisting of 13 items that can be used to measure user trust levels, including competence, reciprocity, and perceived risk. Moreover, these results should not only be limited to HOIA but can also be implemented to measure trust in other CTAs.

Acknowledgments

This study was partly funded by the Trust and Influence Programme (FA8655-22-1-7051), European Office of Aerospace Research and Development, and US Air Force Office of Scientific Research.

Authors' Contributions

Conceptualization: SS, TK; Data collection: TK; Data analysis: SS, TK; Interpretation of results: SS; Writing-original draft: SS.

Conflicts of Interest

None declared.

References

- Du L, Raposo VL, Wang M. COVID-19 contact tracing apps: a technologic tower of Babel and the gap for international pandemic control. *JMIR Mhealth Uhealth* 2020 Nov 27;8(11):e23194. [doi: [10.2196/23194](https://doi.org/10.2196/23194)]
- Williams S, Armitage CJ, Tampe T, Dienes K. Public attitudes towards COVID-19 contact tracing apps: a UK-based focus group study. *Health Expect* 2021 Apr;24(2):377-385. [doi: [10.1111/hex.13179](https://doi.org/10.1111/hex.13179)] [Medline: [33434404](https://pubmed.ncbi.nlm.nih.gov/33434404/)]
- Nguyen T, Miettinen D, Sadeghi AR. Long live randomization: on privacy-preserving contact tracing in pandemic. 2020 Presented at: MTD'20: 7th ACM Workshop on Moving Target Defense at 2020 ACM SIGSAC Conference on Computer and Communications Security; November 9, 2020; virtual. [doi: [10.1145/3411496.3421229](https://doi.org/10.1145/3411496.3421229)]
- Bano M, Zowghi D, Arora C. Requirements, politics, or individualism: what drives the success of COVID-19 contact-tracing apps? *IEEE Softw* 2021 Jan;38(1):7-12. [doi: [10.1109/ms.2020.3029311](https://doi.org/10.1109/ms.2020.3029311)]
- Kainda R, Fléchaïs I, Roscoe AW. Security and usability: analysis and evaluation. 2010 Presented at: 2010 International Conference on Availability, Reliability and Security; February 15-18, 2010; Krakow, Poland p. 275-282. [doi: [10.1109/ares.2010.77](https://doi.org/10.1109/ares.2010.77)]
- Ahmed N, Michelin RA, Xue W, Ruj S, Malaney R, Kanhere SS, et al. A survey of COVID-19 contact tracing apps. *IEEE Access* 2020;8:134577-134601. [doi: [10.1109/access.2020.3010226](https://doi.org/10.1109/access.2020.3010226)]
- Eloff MM, Eloff JHP. Human computer interaction: an information security perspective. In: Ghonaimy MA, El-Hadidi MT, Aslan HK, editors. *Security in the Information Society*. New York: Springer; 2002:535-545.
- King J, Lampinen A, Smolen A. Privacy: is there an app for that? 2011 Presented at: Seventh Symposium On Usable Privacy and Security (SOUPS); July 20-22, 2011; Pittsburgh, PA. [doi: [10.1145/2078827.2078843](https://doi.org/10.1145/2078827.2078843)]

9. ISO/IEC 27000:2009 Information technology — Security techniques — Information security management systems — Overview and vocabulary. International Organization for Standardization. 2009. URL: <https://www.iso.org/standard/41933.html> [accessed 2022-05-19]
10. Ethics guidelines for trustworthy AI. European Commission. 2019 Apr 08. URL: <https://digital-strategy.ec.europa.eu/en/library/ethics-guidelines-trustworthy-ai> [accessed 2022-06-01]
11. Kelton K, Fleischmann KR, Wallace WA. Trust in digital information. *J Am Soc Inf Sci* 2008 Feb 01;59(3):363-374. [doi: [10.1002/asi.20722](https://doi.org/10.1002/asi.20722)]
12. Trivedi A, Vasisht D. Digital contact tracing. *SIGCOMM Comput Commun Rev* 2020 Oct 26;50(4):75-81. [doi: [10.1145/3431832.3431841](https://doi.org/10.1145/3431832.3431841)]
13. COVID-19 teemaline küsitlus 26. Turu-uuringute AS. URL: <https://riigikantselei.ee/media/933/download> [accessed 2021-02-19]
14. Munzert S, Selb P, Gohdes A, Stoetzer LF, Lowe W. Tracking and promoting the usage of a COVID-19 contact tracing app. *Nat Hum Behav* 2021 Feb 21;5(2):247-255. [doi: [10.1038/s41562-020-01044-x](https://doi.org/10.1038/s41562-020-01044-x)] [Medline: [33479505](https://pubmed.ncbi.nlm.nih.gov/33479505/)]
15. Cavoukian A. Privacy by design: The 7 foundational principles. Information and privacy commissioner of Ontario, Canada. 2009. URL: <https://www.ipc.on.ca/wp-content/uploads/resources/7foundationalprinciples.pdf> [accessed 2022-05-19]
16. Chan EY, Saqib NU. Privacy concerns can explain unwillingness to download and use contact tracing apps when COVID-19 concerns are high. *Comput Human Behav* 2021 Jun;119:106718 [FREE Full text] [doi: [10.1016/j.chb.2021.106718](https://doi.org/10.1016/j.chb.2021.106718)] [Medline: [33526957](https://pubmed.ncbi.nlm.nih.gov/33526957/)]
17. Kwasny M, Caine KE, Rogers WA, Fisk AD. Privacy and technology: folk definitions and perspectives. *Proc SIGCHI Conf Hum Factor Comput Syst* 2008;2008:3291-3296 [FREE Full text] [doi: [10.1145/1358628.1358846](https://doi.org/10.1145/1358628.1358846)] [Medline: [29057397](https://pubmed.ncbi.nlm.nih.gov/29057397/)]
18. Simko L, Calo R, Roesner F, Kohno T. COVID-19 contact tracing and privacy: studying opinion and preferences. arXiv. 2020 Dec 17. URL: <https://arxiv.org/abs/2005.06056> [accessed 2022-05-19]
19. Ministry of Social Affairs. COVID-19 blogi. Terviseamet. 2021. URL: <https://www.terviseamet.ee/et/uudised/covid-19-blogi-26-aprill-ooapaevaga-lisandus-163-positiivset-testi> [accessed 2021-04-27]
20. Gulati S, Sousa S, Lamas D. Design, development and evaluation of a human-computer trust scale. *Behav Inf Technol* 2019 Aug 31;38(10):1004-1015. [doi: [10.1080/0144929x.2019.1656779](https://doi.org/10.1080/0144929x.2019.1656779)]
21. Schoorman FD, Mayer RC, Davis JH. An integrative model of organizational trust: past, present, and future. *Acad Manag Rev* 2007 Apr;32(2):344-354. [doi: [10.5465/amr.2007.24348410](https://doi.org/10.5465/amr.2007.24348410)]
22. Sousa S, Dias P, Lamas D. A model for human-computer trust: A key contribution for leveraging trustful interactions. 2014 Presented at: 9th Iberian Conference on Information Systems and Technologies (CISTI); June 18-21, 2014; Barcelona, Spain p. 1-6. [doi: [10.1109/cisti.2014.6876935](https://doi.org/10.1109/cisti.2014.6876935)]
23. Wiedmann KP, Reeh MO, Varelmann D, Hennigs N. The crucial role of user's perceived trust in the orchestration and adoption of IT-ecosystems. 2009 Presented at: MEDES '09: International Conference on Management of Emergent Digital EcoSystems; October 27-30, 2009; France p. 383-390. [doi: [10.1145/1643823.1643894](https://doi.org/10.1145/1643823.1643894)]
24. Egger FN. "Trust me, I'm an online vendor": towards a model of trust for e-commerce system design. 2000 Apr 01 Presented at: CHI EA '00: CHI '00 Extended Abstracts on Human Factors in Computing Systems; April 1-6, 2000; The Hague, the Netherlands p. 101-102. [doi: [10.1145/633292.633352](https://doi.org/10.1145/633292.633352)]
25. Kramer RM. Rethinking trust. *Harvard Business Review*. 2009 Jun. URL: <https://hbr.org/2009/06/rethinking-trust> [accessed 2021-03-19]
26. Pinto A, Sousa S, Silva C, Coelho P. Adaptation and validation of the HCTM Scale into Human-robot interaction Portuguese context: A study of measuring trust in human-robot interactions. 2020 Presented at: 11th Nordic Conference on Human-Computer Interaction: Shaping Experiences, Shaping Society; October 25-29, 2020; Tallinn, Estonia. [doi: [10.1145/3419249.3420087](https://doi.org/10.1145/3419249.3420087)]
27. Magalhães Hill H, Hill A. Investigação por questionário, 2nd edition. Lisbon: Sílabo; 2005:16.
28. Etikan I. Comparison of convenience sampling and purposive sampling. *Am J Theor Appl Stat* 2016;5(1):1. [doi: [10.11648/j.ajtas.20160501.11](https://doi.org/10.11648/j.ajtas.20160501.11)]
29. Hair J, Hult G, Ringle C, Sarstedt M, Sarstedt M. A primer on partial least squares structural equation modeling (PLS-SEM). Thousand Oaks, CA: SAGE Publications, Inc; 2014.
30. Mcknight DH, Carter M, Thatcher JB, Clay PF. Trust in a specific technology. *ACM Trans Manage Inf Syst* 2011 Jun;2(2):1-25. [doi: [10.1145/1985347.1985353](https://doi.org/10.1145/1985347.1985353)]
31. Harwood T, Garry T. Internet of Things: understanding trust in techno-service systems. *J Serv Manag* 2017 Jun 19;28(3):442-475. [doi: [10.1108/JOSM-11-2016-0299](https://doi.org/10.1108/JOSM-11-2016-0299)]
32. Hulland J. Use of partial least squares (PLS) in strategic management research: a review of four recent studies. *Strat Mgmt J* 1999 Feb;20(2):195-204. [doi: [10.1002/\(sici\)1097-0266\(199902\)20:2<195::aid-smj13>3.0.co;2-7](https://doi.org/10.1002/(sici)1097-0266(199902)20:2<195::aid-smj13>3.0.co;2-7)]

Abbreviations

AI: artificial intelligence

AVE: average variance extracted

CTA: contact-tracing app

HCTS: Human Computer Trust Scale

HOIA: Estonian contact-tracing app for COVID-19

PLS-SEM: partial least-squares structural equation modeling

Edited by A Kushniruk; submitted 30.09.21; peer-reviewed by A Tannoubi, L Maaß; comments to author 07.11.21; revised version received 20.02.22; accepted 19.04.22; published 13.06.22.

Please cite as:

Sousa S, Kalju T

Modeling Trust in COVID-19 Contact-Tracing Apps Using the Human-Computer Trust Scale: Online Survey Study

JMIR Hum Factors 2022;9(2):e33951

URL: <https://humanfactors.jmir.org/2022/2/e33951>

doi: [10.2196/33951](https://doi.org/10.2196/33951)

PMID: [35699973](https://pubmed.ncbi.nlm.nih.gov/35699973/)

©Sonia Sousa, Tiina Kalju. Originally published in JMIR Human Factors (<https://humanfactors.jmir.org>), 13.06.2022. This is an open-access article distributed under the terms of the Creative Commons Attribution License (<https://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Human Factors, is properly cited. The complete bibliographic information, a link to the original publication on <https://humanfactors.jmir.org>, as well as this copyright and license information must be included.

Review

Self-care Behaviors and Technology Used During COVID-19: Systematic Review

Fareeya Sakur¹, BCom, MPH; Kaneshia Ward¹, BClinSc, MRes; Neha Nafees Khatri¹, MBBS, MPH; Annie Y S Lau¹, BE, PhD

Australian Institute of Health Innovation, Macquarie University, North Ryde, Australia

Corresponding Author:

Fareeya Sakur, BCom, MPH

Australian Institute of Health Innovation

Macquarie University

75 Talavera Rd

North Ryde, NSW 2113

Australia

Phone: 61 (02) 9850 2400

Email: fareeya.sakur@hdr.mq.edu.au

Abstract

Background: Self-care behaviors are essential for people living with chronic conditions; however, the outbreak of the COVID-19 pandemic has imposed additional complications on their daily routines. Few studies have analyzed how self-care behaviors have changed during COVID-19 and the role of digital technology, especially among people with chronic conditions.

Objective: This study aims to review how self-care behaviors have changed for people with chronic conditions during the COVID-19 pandemic, and what technology they have adopted to manage their conditions during that period.

Methods: A systematic review was conducted using narrative synthesis. Data were extracted from PubMed, MEDLINE, EMBASE, PsycINFO, CINAHL, and Google Scholar, including articles from December 2019 onward. Eligible studies focused on adults diagnosed with chronic conditions undertaking any self-care tasks in line with the middle-range theory of self-care of chronic illness (ie, self-care maintenance, monitoring, and management). The methodological quality of the included articles was assessed with the McMaster Critical Review Forms for Quantitative or Qualitative Studies.

Results: In total, 36 primary research articles were included. Changes to self-care behaviors during COVID-19 among people with chronic conditions were organized according to the middle-range theory of self-care of chronic illness focusing on self-care maintenance (ie, medication adherence, physical activity, and diet control), self-care monitoring (ie, monitoring signs and symptoms), and self-care management (ie, consultations with health care providers). Positive self-care behaviors observed include the following: individuals trying to maintain good glycemic control during COVID-19 increased their medication adherence in 27% (10/36) of studies; and diet control improved in 50% (18/36) of studies. Negative self-care behaviors observed include the following: decline in physical activities and increased sedentariness were observed in 65% (23/36) of studies; poor diet control was observed in 57% (21/36) of studies; and self-monitoring of health status dropped in 43% (15/36) of studies. The use of technology to support self-care of chronic conditions during COVID-19 was reported in 72% (26/36) of studies. The actual use of telehealth in place of physical consultations during COVID-19 was observed in 50% (18/36) of studies, and other digital technologies (eg, social media apps, smartphone apps, web-based platforms, and web browsing) were used in 50% (18/36) of studies. Telehealth was discussed and recommended as the default technology in delivering future health care services during COVID-19 and beyond in 77% (28/36) of studies.

Conclusions: This review highlighted the necessity to rethink how models of self-care should continue to address the demands of chronic conditions while being responsive to the imminent threats of infectious diseases. Perhaps the silver lining of COVID-19 is that adoption of digital technology (especially telehealth) among a vast cross-section of people with chronic conditions is possible. Future research should investigate effective ways to incorporate evidence-based digital health tools into these new models of self-care that address the challenges of chronic and infectious conditions.

(*JMIR Hum Factors* 2022;9(2):e35173) doi:[10.2196/35173](https://doi.org/10.2196/35173)

KEYWORDS

self-care; self-management; chronic conditions; COVID-19; pandemic; technology; digital health; telehealth; health technology

Introduction

Background

Chronic conditions are an ongoing cause of disability, ill health, and premature death worldwide, and the World Health Organization defines chronic conditions as conditions that are noncommunicable and are of long duration and slow progression [1]. Self-care is essential for people with chronic conditions in order to maintain good control of their health [2]. People with chronic conditions need continuity of care to ensure their conditions are well maintained [3]. However, the public health response to the pandemic (eg, lockdown measures and social distancing) has significantly disrupted this continuity and thus affected people with chronic conditions [4-7].

People with major chronic conditions are not only at a higher risk of COVID-19 infection, but also of worsening their chronic disease outcomes during the pandemic [6-8]. Clinical studies in the United States and Italy undertaken on patients with COVID-19 found that the severity rates and death rates among patients with underlying chronic conditions were 7 times higher than those of patients with nonchronic conditions [8]. From an individual perspective, self-care behaviors have been significantly affected during the pandemic [4,5]. People have had their in-person health care appointments converted to teleconsultations [9]. Many have experienced disruptions in their medication supplies [10]; had limited access to investigative tests (eg, blood tests) [11]; were confronted with barriers to physical activities (PA) [12], imbalanced diets (eg, disruptions in access to food sources) [13], as well as disrupted routines and supplies to necessities; and many have experienced social isolation (eg, not being able to see family and friends), anxiety, and mental distress [12].

The impact of the pandemic on self-care behaviors of people living with chronic conditions was assessed in 2 rapid reviews conducted by Kendzerska et al [12] and Hartmann-Boyce and Mahtani [14] in 2020.

To our knowledge, systematic review-based evidence on how chronic disease self-care behaviors have changed during COVID-19, and how people with chronic conditions have adopted the use of technological aids during COVID-19 to sustain their self-care behaviors remains lacking.

Objectives

In this study, we will undertake a systematic review to examine how self-care behaviors among people with chronic conditions have changed during COVID-19, and the role of digital technology in facilitating those changes. The research questions in this systematic review are as follows: “How have self-care behaviors among people with chronic conditions changed as a result of COVID-19” and “What technological aids have people with chronic conditions used (or adopted) for self-care during COVID-19?”

Methods

Materials and Methods

This systematic review was registered in the International Prospective Register of Systematic Reviews (PROSPERO) with the registration number CRD42021274000.

The review is in compliance with the PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) statement [15]. Details of the PRISMA checklist can be found in [Multimedia Appendix 1](#).

Search Strategy

A modified population, interventions, comparisons, and outcomes (“PICO”) strategy was used to search, with “Population” corresponding to “people with chronic conditions” and “Intervention” as the “self-management of chronic conditions (and the use of technological aids) during COVID-19”; “Comparison” is described as “self-management (and the use of technological aids) before COVID-19,” and “Outcomes” are “changes in self-care behaviors and the use of technology.”

A search from March 6, 2021, to March 11, 2021, was conducted in PubMed, MEDLINE, EMBASE, PsycINFO, CINAHL, and Google Scholar, including all articles published from December 2019. Search terms were designed to capture publications on people living with chronic conditions, their self-care behaviors during COVID-19, and any use of technological aids. [Multimedia Appendix 2](#) provides the complete search strategy.

Inclusion and Exclusion Criteria

Articles were eligible if they had the following criteria: focused on adults diagnosed with chronic conditions (conditions that limits self-care, requires medical interventions, and lasts more than 6 months); included a quantitative or qualitative component; focused on reporting self-care tasks during COVID-19, undertaken by people diagnosed with chronic conditions; included use of technology in self-care of chronic conditions; and were published in the English language from December 2019.

Articles were excluded if they had the following: did not focus on people with chronic conditions (eg, caregivers or care providers); were not COVID-19-related; focused on purely educational programs to improve self-management of chronic conditions; focused on technology only with no outcome measures; and were protocol papers or opinion articles.

[Multimedia Appendix 3](#) provides the complete criteria.

Study Screening

Full details on abstracts, full-text screening, and data extraction are provided in [Multimedia Appendix 4](#). Each abstract was screened independently by 3 reviewers, disagreements were resolved by consensus, and full-text screening was undertaken by 1 reviewer.

Data extraction was led by 1 reviewer, and a narrative synthesis was conducted to synthesize the findings of the studies. The 36 included articles were read in full, and specific details on self-care behaviors were extracted and organized into the themes of physical control, medication adherence, diet control, monitoring health status, and consultations with health care providers in a tabular form. Specific items on use and recommendation of technology were extracted and summarized in a tabular form and presented in appendices.

Methodological Quality Assessment

The McMaster Critical Appraisal Tools for Quantitative Studies and Qualitative Studies was used [16]. Each individual component is rated as “yes,” “no,” “not addressed,” or “not applicable.” A score of 1 was given to “yes,” 0 to “no” and “not addressed,” while items rated as “not applicable” were removed from the total score. Quantitative studies were assessed over 8 main components of study purpose, literature review, study design, sample, outcomes, intervention, results, and conclusions—with the maximum total score being 14. Qualitative studies were assessed over 8 components, which

were study purpose, literature review, study design, sampling, data collection, data analysis, overall rigor, and conclusions—with the maximum total score being 22. Methodological quality score rating did not warrant exclusion of studies. The results of the assessment of methodological quality are outlined in [Multimedia Appendix 5](#).

Theoretical Framework

The changes in self-care during COVID-19 and the technology used by people living with chronic conditions were reported according to the middle-range theory of self-care of chronic illness. This theory arose from clinical experience caring for persons with heart failure in 2012 [2]. Self-care is described as the maintenance of health. It is a process undertaken through health promotion practices and management of health conditions that can be performed in a healthy or ill state [2]. The focus is on the following three key concepts: self-care maintenance, self-care monitoring, and self-care management ([Figure 1](#)) [2]. The operational definitions and examples of the three key concepts are outlined in [Table 1](#) [2,17].

Figure 1. Middle-range theory of self-care of chronic illness model based on three key concepts of self-care behavior [2].



Table 1. The operational definitions and examples of the three key concepts of the middle-range theory of self-care for chronic illness.

Self-care behaviors	Definition	Examples
Self-care maintenance	Behaviors adopted by people to maintain physical and emotional stability for their chronic conditions. They can be strategies set by the individuals alone or based upon recommendations determined between individuals and their health care providers. These behaviors can be related to lifestyle such as diet control, physical exercise, and taking prescribed medications [2,17].	Combining 15 min of postmeal walking with 30 min of resistance training.
Self-care monitoring	A process that involves routinely observing for changes in signs and symptoms with vigilance and acting accordingly [2,17]. It encompasses systematic and routine monitoring. Individuals that are skilled in monitoring their symptoms and communicating them to their health support team help produce the best health outcomes [2,17].	Checking their blood glucose levels daily.
Self-care management	Evaluating changes in signs and symptoms (from both emotional and physical well-being perspectives) that are present due to sickness, treatments undertaken, or the environment. If a response is needed, then a treatment plan can be sought, implemented, and evaluated. The efficacy of the treatment plan in achieving the desired health outcomes is assessed on an ongoing basis, between the individual and their health care team [2,17].	During monitoring, if blood glucose levels are elevated, then a treatment plan can be set in consultation with their health care provider.

Results

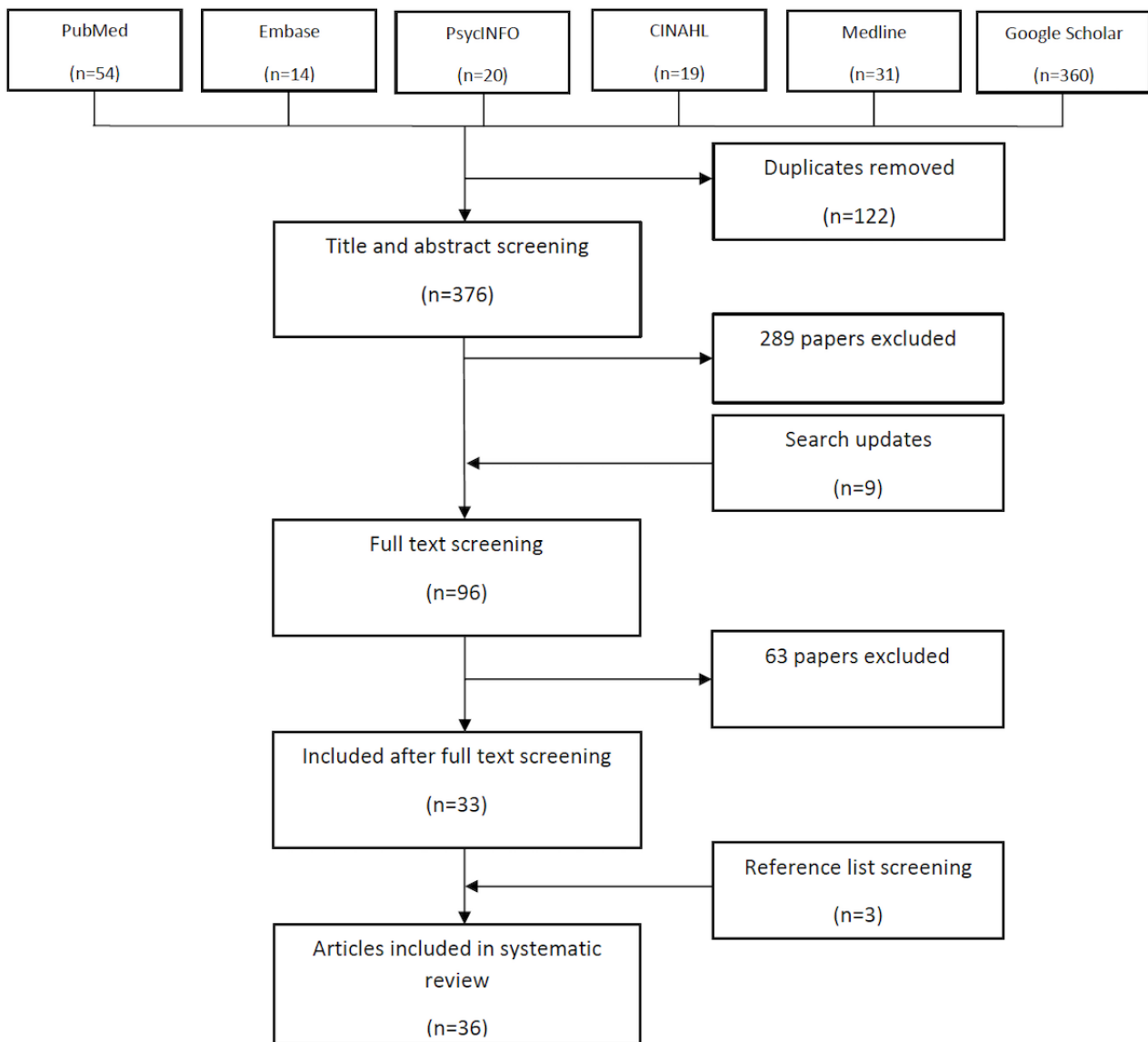
Screening Process

The database search retrieved 498 publications, and 122 duplicates were removed. After title and abstract screening, 289 publications were removed. Search updates led to 9 publications being included in the screening. After full text screening, 63

publications were excluded, leaving 33 included articles. A further 3 articles were identified by searching the reference lists of the included articles. The entire screening process concluded with the inclusion of 36 original research publications.

The literature selection process is outlined in [Figure 2](#). [Multimedia Appendix 4](#) provides more details about the screening process.

Figure 2. Number of articles included at each stage of the screening process.



Characteristics of the Included Articles

A total of 36 original research publications were included in the systematic review. The most common research strategy employed was semistructured online interviews. The majority of studies were conducted in India (10/36, 28%), followed by United Kingdom (5/36, 14%); United States and Brazil (4/36, 11%); Spain (3/36, 8%); Hong Kong and Arab states (2/36, 6%); and Pakistan, Israel, Denmark, Poland, Canada, Saudi Arabia, and Australia (1/36, 3%).

Over 28 health conditions were studied in the included articles: diabetes mellitus (26/36, 72%), hypertension (8/36, 22%), cancer (7/36, 19%), cardiovascular diseases (6/36, 17%), chronic heart diseases (6/36, 17%), respiratory conditions (6/36, 17%), chronic kidney diseases (5/36, 17%), and chronic obstructive pulmonary disease (4/36, 11%).

The different health conditions included in the articles are further outlined in [Multimedia Appendix 6](#). The conditions most frequently stated in the included studies are diabetes mellitus (26/36, 72%), hypertension (8/36, 22%), and cancer (7/36, 19%).

The self-care behaviors included in the articles are provided in [Multimedia Appendix 7](#). Instances where technology was stated in the included articles are outlined in [Multimedia Appendix 8](#). Additionally, [Multimedia Appendices 9](#) and [10](#) provide more information on the included articles.

Self-care Behaviors Adopted by Participants

People living with chronic conditions are embodied with the responsibility for their daily care and actively engage in tasks essential for long-term management of their conditions [2,17].

Self-care behaviors observed in the included studies are organized according to the middle-range theory of self-care of chronic illness. These include self-care maintenance (ie, medication adherence, physical activity, and diet control), self-care monitoring (ie, monitoring signs and symptoms), and self-care management (ie, consultations with health care providers). The technology reported in these studies was organized according to the technology used for self-care during COVID-19 and technology recommended for self-care during and beyond COVID-19. The results of self-care behaviors found in the included articles are illustrated in [Table 2](#).

Table 2. Results of self-care behaviors in the included articles (N=36).

Self-care behaviors	Increase reported in the studies, n (%)	Decrease reported in the studies, n (%)
Self-care maintenance		
Medication adherence (n=15)	4 (27)	8 (53)
Physical activities (n=20)	7 (35)	13 (65)
Diet control (n=14)	7 (50)	8 (57)
Self-care monitoring		
Self-monitoring of signs and symptoms (n=7)	4 (57)	3 (43)
Self-care management		
Consultations with health care providers (n=23)	13 ^a (57)	23 (100)

^aReplaced with telehealth.

Expanded results of self-care behaviors and technology adopted or recommended in the included articles are outlined in [Multimedia Appendix 11](#).

Self-care Maintenance

Medication Adherence

A total of 15 studies reported on medication adherence during COVID-19 [18-32]. Increase in medication adherence among participants trying to maintain good glycemic control during the pandemic was reported in 27% (4/15) [18,19,23,24] of the studies. Participants had difficulties in sourcing medication due to supply shortages in 53% (8/15) of these studies [20-22,25,27,29,30,32], with people in rural areas faring worse [29].

In one study, the participants reported lower levels of medication adherence due to store closures, fears of contracting the virus if they went outside, and difficulties in purchasing caused by financial constraints due to job losses [30]. A few studies [25,28,32] reported on difficulties in obtaining prescription renewals among participants due to cancellation of physical consultations, and telehealth was used for prescription renewals.

PA

A total of 20 studies reported on PA during the pandemic [18,20,24,26-29,33-45], and 35% (7/20) of these studies [28,33,38,40-43] reported on continual or increase in PA among the participants in their studies. Disruptions in routines, fear of going outdoors, lack of motivation, as well as increased anxiety and stress during COVID-19 led to the decline in PA in 65% (13/20) of these studies [18,20,24,26,27,29,34-37,39,44,45].

The participants in a study that measured PA exclusively [34] reported that 71.5% of Brazilian adults were not meeting the minimum PA recommendations. The study revealed that increasing age and multimorbidity had a positive association with increased sedentary risk during the pandemic.

Diet Control

A total of 14 studies reported on dietary habits during the pandemic [18-20,24,27,29,30,33,35,37,38,41,45,46], and 50% (7/14) of these studies [19,20,33,37,38,41,46] reported improved diet management among their participants. Reasons were of increased time availability, with more meals being prepared at home and lockdown restrictions limiting takeaway consumptions among participants.

There was an association between better diet control and improved glycemic control reported among participants [19,33,37,38,41,46]. Poor diet control was found in 57% (8/14) of these studies [18,20,24,27,29,30,35,45], with participants reducing their consumption of fresh fruit and vegetables due to unavailability, as there were with disruptions to supplies and reduced frequency in shopping to minimize exposure during the pandemic [25,28,32]. Increase in unhealthy food consumption was due to multiple factors such as more sedentary time at home and changes in mood including lack of motivation, boredom, increased anxiety, and stress.

Self-care Monitoring: Self-monitoring of Signs and Symptoms

In total, 7 studies reported on participants monitoring their diabetes mellitus [18,19,24,33,37,38,41]. Regular compliance or increase in monitoring of blood glucose levels during the pandemic among participants was found in 57% (4/7) of these

studies [18,19,24,37]. Decline in monitoring of blood glucose levels among participants were due to difficulties in sourcing testing strips and lack of knowledge barriers in 43% (3/7) of studies [33,38,41].

Self-care Management: Consultations With Health Care Providers

All 23 studies that examined access to health care providers during the pandemic found disruptions to health care services, with postponement or cancellation of consultations noted among

their participants [19-25,27-32,40,42,44,45,47-52]. In the 23 studies that reported on access to health care providers, 57% (13/23) of these studies [19,20,23,24,29,31,32,36,42,46,49,52,53] revealed participants used telehealth services in place of physical consultations with their health care team. Moreover, 13% (3/23) of these studies [21,40,45] found that difficulties in accessing health care services during the pandemic led to issues with glycemic control among the individuals. The results of technology used in the included studies are illustrated in [Table 3](#).

Table 3. Use of technology for self-care of chronic conditions in the included studies (n=26).

Technology reported in the studies	Values, n (%)
Telehealth used during COVID-19 among participants	13 (50)
Other digital technology (television, social media apps, smartphone apps, web-based digital health tools, web-based platforms, and web browsing)	13 (50)
Role of telehealth discussed and recommended	20 (77)

Technology: Technological Aids Used

In total, 26 studies discussed the role of technology during COVID-19 to support individuals' self-care of chronic conditions [19-29,31,32,35,36,40,42,43,45,46,48-53], and 50% (13/26) of the studies [19,20,23,24,29,31,32,36,42,46,49,52,53] reported on the use of telehealth, due to in-person consultations having been replaced with telephone or video consultations. Participants used telehealth for prescription renewals, test results discussion, or simple follow-ups. Moreover, 50% (13/26) of the participants in these studies [19,20,23,24,29,31,32,36,42,46,49,52,53] expressed that telehealth allowed continuity of care for them during the pandemic, that the support helped them maintain their self-care behaviors, and that they would continue using it in the future.

The use of television, social media apps, smartphone apps, web-based digital health tools, web-based platforms, and web browsing was found in 50% (13/26) of these studies [19,20,26-29,32,35,36,43,48,52,53]. One study [36] reported people living with diabetes and liver disease were the highest users of social media, while video consultations were mostly used by people living with chronic liver diseases and neurological conditions in their population sample.

The role of telehealth was discussed and recommended in the future delivery of health care services in 77% (20/26) of these studies [19-22,24-26,29,31,32,36,40,42,45,46,49-53], especially for people living in rural areas [46,50]. According to one study [48], effective intervention strategies are needed to improve digital literacy among elderly people living with chronic conditions to facilitate their participation and presence in digital health.

Telehealth was the most used technology, followed by social media apps (Facebook), messaging apps (WhatsApp, Messenger, and WeChat), web-based platforms for education and exercise (YouTube and web-based exercise platforms), and web browsing (Google).

[Multimedia Appendix 12](#) provides details on the types of technology used and recommended in the included articles.

Discussion

Principal Findings

To our knowledge, this is the first systematic review of changes in self-care behaviors in people with chronic conditions and the technological aids they adopted in managing their conditions during COVID-19.

The purpose of this systematic review was to analyze the existing literature on how self-care behaviors have changed during COVID-19, and the range of technology adopted by people with chronic conditions in managing their conditions during the pandemic. Our results indicate that the lives of people with chronic conditions were altered by the course of measures imposed to contain the spread of COVID-19, with disruptions to their daily routines challenging their self-care behaviors. The lockdown resulted in both favorable and unfavorable changes in self-care behaviors, which could have short- and long-term effects on health.

Positive self-care behaviors that resulted from the lockdown were found among individuals motivated to keep good glycemic control, and those who maintained or increased their medication adherence during COVID-19. Improved diet control resulted from an increase in home cooking and less consumption of takeaways. Cancelled physical consultations were replaced with telehealth to allow continuity of health care services.

Negative self-care behaviors that resulted from the lockdown were from fluctuations in medical supplies, difficulties in sourcing prescriptions, and financial constraints impacting medication adherence. Reduction in fresh produce consumption due to supply issues and lack of motivation led to poor diet control. Significant decline in PA and increased sedentariness were found among participants in most studies during lockdown. In-person visits for routine consultations were postponed or cancelled. Access to health care services was facilitated by telehealth through phone or video consultations to allow continuity of care during the pandemic. However, in some developing countries, proactive contact with people with chronic

diseases during the pandemic with telehealth was rare [30,33,34,41,44,47].

The role of technology in the home setting to manage chronic conditions remains low with telehealth being the most frequently used technology during COVID-19, followed by internet browsing, social media platforms, and messaging apps. There is a lack of studies focusing on the effects of eHealth, mobile health, and health apps in the delivery of health care services or management of self-care during COVID-19; this then presents an opportunity for future research in this area.

Strengths and Limitations

This review has several strengths. We developed and followed a rigorous and predefined protocol that was registered with the International Prospective Register of Systematic Reviews (PROSPERO) database at the beginning of the study. To ensure sensitivity and specificity, we developed an extensive search strategy of literature with the help of a clinical librarian. Eligibility criteria were objectively stated and applied in the screening of each article by 3 independent reviewers, and there was substantial agreement with the full text screening results.

There were some limitations in the review, as only articles published in English were included, and we did not have access to studies in other languages. The use of validated instruments to measure the effect of COVID-19 public health measures on self-care behaviors across all studies was limited. The review focused only on self-care behaviors undertaken by people living with chronic conditions, leaving out the caregiver's role in managing self-care. Only 4 studies used qualitative approaches, and there is a need to increase the use of qualitative methodology in self-care research to gain more insights or context on the circumstances involved.

It is important to note that search of databases consisted of keywords such as "chronic conditions" and "multimorbidity" and not the exact diagnosis terms, which may result in excluding articles that use exact diagnosis terms (eg, "diabetes" and "cardiovascular"). The majority of studies examined technology that will enable communications with people with chronic conditions and their health care provider. However, the use of other digital tools that help in monitoring and providing aid in managing their conditions was limited. There is a need for further research on the use of other types of technology and how it was used in the management of self-care in the home environment.

Comparison With Existing Literature

A recent review focusing on lifestyle changes during COVID-19 [54] found increased consumption of unhealthy food and decline in PA across various population groups. The use of telehealth was widespread, and the review proposed the use of virtual networks in the future delivery of health care services, which is in line with our findings.

Kendzerska et al [12] focused on chronic disease management in the primary and specialty care settings. There were concerns that medicine shortages during COVID-19 and the decline in physical activity found among people with chronic conditions could exacerbate their conditions. The implementation of

telemedicine during the pandemic outbreak has been associated with many barriers especially among elderly patients with digital literacy being a common issue, as we observed in our included study [12].

Most reviews on people with chronic conditions during the pandemic analyzed the prevalence of chronic conditions in patients diagnosed with COVID-19 and the adverse clinical outcomes associated with the population group. Increased age and underlying chronic conditions were the strongest predictors of longer hospitalizations or mortality rates among patients diagnosed with COVID-19 [55-57].

In our review, the lived experiences of people with chronic conditions and how they managed their self-care behaviors during the pandemic were examined; we also analyzed the role of digital technology in facilitating them.

Implications

The COVID-19 experience provides an opportunity to rethink what worked and what did not during the pandemic, and to better prepare for future pandemics or health threats.

Key Implications 1: Self-care Behaviors

The evidence from this review shows that certain groups of people with chronic conditions managed to improve or continue with their self-care behaviors amid the pandemic while others struggled to manage them. There is a need for in-depth study on how certain population groups were able to maintain this behavior and the coping strategies they adopted.

These findings can be drawn upon to enhance current self-care interventions to further empower and support these individuals in sustaining their self-care capabilities. It can help individuals to independently cope with self-care behaviors and maintain positive health outcomes, particularly in circumstances when health care resources are redirected toward infectious disease control.

Further research is required on why certain individuals failed to engage in effective self-care behaviors during COVID-19. The factors or barriers that affected this adverse behavior needs to be investigated. The findings can be used to develop successful strategies or interventions to reinforce better health-promoting behaviors and increased adherence to self-care behaviors among this population group. The experiences of COVID-19 have shown how integral self-care is in chronic conditions management. The health care system should use this opportunity to work on a systemic approach to tackle health inequities and incorporate self-care management into the fabric of health care services. Health care professionals also need to evaluate how able individuals are in understanding the information on self-care behaviors provided to them, and their capability to engage in self-care independently.

Hence, health care professionals should tailor self-care advice and plan at the individuals' level of understanding, their capacity, and the context they are in, so that their actions are effective and sustainable for a longer period of time.

Key Implications 2: Digital Technology Adoption

The COVID-19 outbreak has changed the conversation on digital interventions in health care services. The rapid adoption of telehealth, as well as the tidal of acceptance by individuals and health care providers in the delivery of health care services, has led to telehealth emerging as the silver lining of the pandemic. It has re-envisioned chronic care management and opened opportunities of using evidence-based digital health interventions that can promote and support self-care capabilities among people with chronic conditions both now and in any future public health crisis.

The reassignment of the health care resources during COVID-19 on prioritizing communicable disease care severely disrupted chronic care management with cancelled or postponed health care services. This led to a backlog of routine services and a decline in screening and preventive care that could later exacerbate health risks and strain the health care system. Care pathways need to be reconfigured to allow new models of health care to treat both communicable and chronic diseases continuously. Embedding and accelerating digital changes in chronic care management can instigate individuals and health care providers to work on solutions that allow chronic care management to be maintained alongside communicable diseases in future pandemics or health threats. Future digital health interventions should consider the influence of family and friends in the health management of people with chronic condition. They play a major role in supporting or assisting individuals with making daily decisions about medications and symptoms management, helping coordinate health care services and facilitate healthy behavior changes. It is important that newer digital health interventions recognize and provide digital solutions for all members of the individual care team for optimal health outcomes.

There is a lack of research on vulnerable population groups (ie, elderly, indigenous communities, and disability groups) who are at a greater risk of contracting COVID-19 and the associated

population health implications. These population groups must be the focus of future studies, evaluating their lived experiences in managing their chronic conditions and use of technology during COVID-19. Any disparities identified in access, digital literacy, and equity should be appropriately addressed. The lessons of the pandemic should not be lost; they should be used to build new approaches in chronic self-management.

Conclusions

This review provided insights into how people with chronic conditions managed their self-care behaviors during COVID-19, and the types of technology used during that period. In our systematic review, we found that the measures imposed to mitigate the spread of COVID-19 did have an impact on people with chronic conditions and their self-care capabilities, resulting in the decline in PA and self-monitoring of signs and symptoms, increase in unhealthy food consumption, and difficulties in medication adherence.

There are concerns that if these negative self-care behaviors are sustained postpandemic, they could lead to further health complications among people with underlying chronic conditions and burden the overstretched health care system.

The lived experiences of COVID-19 should become a catalyst for adoption of a new model for health care that is flexible to respond to both chronic and infectious diseases. It should recognize and have measures in place to support and enhance self-care capabilities among people with underlying chronic conditions during the pandemic, and for future health threats. The use of digital technology (telehealth, online platforms, and messaging apps) connected individuals to health care services and changed the way they receive care during the pandemic. This highlights the need for further research on incorporating and leveraging evidence-based digital health tools into newer models of health care. These can then aim to engage and motivate individuals toward the effective management of their self-care behaviors and facilitate continuity of health care services in any situation.

Acknowledgments

The authors would like to thank Urvashi Rohilla for her assistance with the paper screening.

Authors' Contributions

FS designed the study, developed the search strategy, conducted the searches, screened the retrieved papers, extracted relevant information, and drafted the paper as the first author. KW and NNK contributed toward screening. AYSL contributed throughout the paper starting from conceptualization to editing the subsequent drafts of the paper.

Conflicts of Interest

None declared.

Multimedia Appendix 1

PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) 2020 checklist.

[[DOCX File , 1268 KB](#) - [humanfactors_v9i2e35173_app1.docx](#)]

Multimedia Appendix 2

Search strategy.

[[DOCX File , 426 KB - humanfactors_v9i2e35173_app2.docx](#)]

Multimedia Appendix 3

Inclusion and exclusion criteria.

[[DOCX File , 15 KB - humanfactors_v9i2e35173_app3.docx](#)]

Multimedia Appendix 4

Search and screening process.

[[DOCX File , 14 KB - humanfactors_v9i2e35173_app4.docx](#)]

Multimedia Appendix 5

Methodological quality assessment.

[[DOCX File , 542 KB - humanfactors_v9i2e35173_app5.docx](#)]

Multimedia Appendix 6

Health conditions stated in the included studies.

[[DOCX File , 14 KB - humanfactors_v9i2e35173_app6.docx](#)]

Multimedia Appendix 7

Self-care behaviors included in the studies.

[[DOCX File , 14 KB - humanfactors_v9i2e35173_app7.docx](#)]

Multimedia Appendix 8

Technology stated in the included studies.

[[DOCX File , 13 KB - humanfactors_v9i2e35173_app8.docx](#)]

Multimedia Appendix 9

Characteristics of the included studies.

[[DOCX File , 24 KB - humanfactors_v9i2e35173_app9.docx](#)]

Multimedia Appendix 10

Characteristics of the included studies.

[[DOCX File , 21 KB - humanfactors_v9i2e35173_app10.docx](#)]

Multimedia Appendix 11

Results.

[[DOCX File , 18 KB - humanfactors_v9i2e35173_app11.docx](#)]

Multimedia Appendix 12

Technology used and recommended.

[[DOCX File , 22 KB - humanfactors_v9i2e35173_app12.docx](#)]

References

1. Bernell S, Howard SW. Use Your Words Carefully: What Is a Chronic Disease? *Front Public Health* 2016 Aug 02;4:159 [[FREE Full text](#)] [doi: [10.3389/fpubh.2016.00159](https://doi.org/10.3389/fpubh.2016.00159)] [Medline: [27532034](https://pubmed.ncbi.nlm.nih.gov/27532034/)]
2. Riegel B, Jaarsma T, Strömberg A. A middle-range theory of self-care of chronic illness. *Advances in Nursing Science* 2012;35(3):194-204. [doi: [10.1097/ans.0b013e318261b1ba](https://doi.org/10.1097/ans.0b013e318261b1ba)]
3. Harris MF, Zwar NA. Care of patients with chronic disease: the challenge for general practice. *Med J Aust* 2007 Jul 16;187(2):104-107. [doi: [10.5694/j.1326-5377.2007.tb01152.x](https://doi.org/10.5694/j.1326-5377.2007.tb01152.x)] [Medline: [17635094](https://pubmed.ncbi.nlm.nih.gov/17635094/)]
4. Bavli I, Sutton B, Galea S. Harms of public health interventions against covid-19 must not be ignored. *BMJ* 2020 Nov 02;371:m4074. [doi: [10.1136/bmj.m4074](https://doi.org/10.1136/bmj.m4074)] [Medline: [33139247](https://pubmed.ncbi.nlm.nih.gov/33139247/)]
5. Levene LS, Seidu S, Greenhalgh T, Khunti K. Pandemic threatens primary care for long term conditions. *BMJ* 2020 Oct 05;371:m3793. [doi: [10.1136/bmj.m3793](https://doi.org/10.1136/bmj.m3793)] [Medline: [33020049](https://pubmed.ncbi.nlm.nih.gov/33020049/)]
6. Danhieux K, Buffel V, Pairon A, Benkheil A, Remmen R, Wouters E, et al. The impact of COVID-19 on chronic care according to providers: a qualitative study among primary care practices in Belgium. *BMC Fam Pract* 2020 Dec 05;21(1):255 [[FREE Full text](#)] [doi: [10.1186/s12875-020-01326-3](https://doi.org/10.1186/s12875-020-01326-3)] [Medline: [33278877](https://pubmed.ncbi.nlm.nih.gov/33278877/)]

7. Flaherty GT, Hession P, Liew CH, Lim BCW, Leong TK, Lim V, et al. COVID-19 in adult patients with pre-existing chronic cardiac, respiratory and metabolic disease: a critical literature review with clinical recommendations. *Trop Dis Travel Med Vaccines* 2020 Aug 28;6(1):16 [FREE Full text] [doi: [10.1186/s40794-020-00118-y](https://doi.org/10.1186/s40794-020-00118-y)] [Medline: [32868984](https://pubmed.ncbi.nlm.nih.gov/32868984/)]
8. Xia Y, Li Q, Li W, Shen H. Elevated mortality of chronic diseases during COVID-19 pandemic: a cause for concern? *Ther Adv Chronic Dis* 2020 Sep 26;11:2040622320961590 [FREE Full text] [doi: [10.1177/2040622320961590](https://doi.org/10.1177/2040622320961590)] [Medline: [33062237](https://pubmed.ncbi.nlm.nih.gov/33062237/)]
9. Isautier JM, Copp T, Ayre J, Cvejic E, Meyerowitz-Katz G, Batcup C, et al. People's Experiences and Satisfaction With Telehealth During the COVID-19 Pandemic in Australia: Cross-Sectional Survey Study. *J Med Internet Res* 2020 Dec 10;22(12):e24531 [FREE Full text] [doi: [10.2196/24531](https://doi.org/10.2196/24531)] [Medline: [33156806](https://pubmed.ncbi.nlm.nih.gov/33156806/)]
10. Choo EK, Rajkumar SV. Medication Shortages During the COVID-19 Crisis: What We Must Do. *Mayo Clin Proc* 2020 Jun;95(6):1112-1115 [FREE Full text] [doi: [10.1016/j.mayocp.2020.04.001](https://doi.org/10.1016/j.mayocp.2020.04.001)] [Medline: [32312491](https://pubmed.ncbi.nlm.nih.gov/32312491/)]
11. Verhoeven V, Tsakitzidis G, Philips H, Van Royen P. Impact of the COVID-19 pandemic on the core functions of primary care: will the cure be worse than the disease? A qualitative interview study in Flemish GPs. *BMJ Open* 2020 Jun 17;10(6):e039674 [FREE Full text] [doi: [10.1136/bmjopen-2020-039674](https://doi.org/10.1136/bmjopen-2020-039674)] [Medline: [32554730](https://pubmed.ncbi.nlm.nih.gov/32554730/)]
12. Kendzerska T, Zhu DT, Gershon AS, Edwards JD, Peixoto C, Robillard R, et al. The Effects of the Health System Response to the COVID-19 Pandemic on Chronic Disease Management: A Narrative Review. *RMHP* 2021 Feb;Volume 14:575-584. [doi: [10.2147/rmhp.s293471](https://doi.org/10.2147/rmhp.s293471)]
13. Wicaksana AL, Hertanti NS, Ferdiana A, Pramono RB. Diabetes management and specific considerations for patients with diabetes during coronavirus diseases pandemic: A scoping review. *Diabetes Metab Syndr* 2020 Sep;14(5):1109-1120 [FREE Full text] [doi: [10.1016/j.dsx.2020.06.070](https://doi.org/10.1016/j.dsx.2020.06.070)] [Medline: [32659694](https://pubmed.ncbi.nlm.nih.gov/32659694/)]
14. Hartmann-Boyce J, Mahtani KR. Supporting people with long-term conditions (LTCs) during national emergencies. The Centre for Evidence-Based Medicine. 2020 Mar 25. URL: <https://www.cebm.net/covid-19/supporting-people-with-long-term-conditions-ltcs-during-national-emergencies/> [accessed 2022-05-09]
15. Moher D, Liberati A, Tetzlaff J, Altman DG, PRISMA Group. Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. *Int J Surg* 2010;8(5):336-341 [FREE Full text] [doi: [10.1016/j.ijisu.2010.02.007](https://doi.org/10.1016/j.ijisu.2010.02.007)] [Medline: [20171303](https://pubmed.ncbi.nlm.nih.gov/20171303/)]
16. Orlando JF, Beard M, Kumar S. Systematic review of patient and caregivers' satisfaction with telehealth videoconferencing as a mode of service delivery in managing patients' health. *PLoS One* 2019 Aug 30;14(8):e0221848 [FREE Full text] [doi: [10.1371/journal.pone.0221848](https://doi.org/10.1371/journal.pone.0221848)] [Medline: [31469865](https://pubmed.ncbi.nlm.nih.gov/31469865/)]
17. Gupta S, Lakshmi PM, Kaur M, Rastogi A. Role of self-care in COVID-19 pandemic for people living with comorbidities of diabetes and hypertension. *J Family Med Prim Care* 2020 Nov;9(11):5495-5501 [FREE Full text] [doi: [10.4103/jfmpc.jfmpc_1684_20](https://doi.org/10.4103/jfmpc.jfmpc_1684_20)] [Medline: [33532385](https://pubmed.ncbi.nlm.nih.gov/33532385/)]
18. Alshareef R, Al Zahrani A, Alzaharani A, Ghandoura L. Impact of the COVID-19 lockdown on diabetes patients in Jeddah, Saudi Arabia. *Diabetes Metab Syndr* 2020 Sep;14(5):1583-1587 [FREE Full text] [doi: [10.1016/j.dsx.2020.07.051](https://doi.org/10.1016/j.dsx.2020.07.051)] [Medline: [32947759](https://pubmed.ncbi.nlm.nih.gov/32947759/)]
19. Anjana RM, Pradeepa R, Deepa M, Jebarani S, Venkatesan U, Parvathi SJ, et al. Acceptability and Utilization of Newer Technologies and Effects on Glycemic Control in Type 2 Diabetes: Lessons Learned from Lockdown. *Diabetes Technol Ther* 2020 Jul 01;22(7):527-534. [doi: [10.1089/dia.2020.0240](https://doi.org/10.1089/dia.2020.0240)] [Medline: [32522031](https://pubmed.ncbi.nlm.nih.gov/32522031/)]
20. Singh K, Kaushik A, Johnson L, Jaganathan S, Jarhyan P, Deepa M, et al. Patient experiences and perceptions of chronic disease care during the COVID-19 pandemic in India: a qualitative study. *BMJ Open* 2021 Jun 18;11(6):e048926 [FREE Full text] [doi: [10.1136/bmjopen-2021-048926](https://doi.org/10.1136/bmjopen-2021-048926)] [Medline: [34145019](https://pubmed.ncbi.nlm.nih.gov/34145019/)]
21. Ugliara Barone MT, Harnik SB, Chaluppe M, Vieira de Luca P, Ngongo B, Pedrosa HC, et al. Decentralized COVID-19 measures in Brazil were ineffective to protect people with diabetes. *Diabetes Metab Syndr* 2020 Nov;14(6):1973-1978 [FREE Full text] [doi: [10.1016/j.dsx.2020.10.005](https://doi.org/10.1016/j.dsx.2020.10.005)] [Medline: [33075740](https://pubmed.ncbi.nlm.nih.gov/33075740/)]
22. Chan EYY, Kim JH, Lo ESK, Huang Z, Hung H, Hung KKC, et al. What Happened to People with Non-Communicable Diseases during COVID-19: Implications of H-EDRM Policies. *Int J Environ Res Public Health* 2020 Aug 03;17(15):5588 [FREE Full text] [doi: [10.3390/ijerph17155588](https://doi.org/10.3390/ijerph17155588)] [Medline: [32756382](https://pubmed.ncbi.nlm.nih.gov/32756382/)]
23. Glintborg B, Jensen DV, Engel S, Terslev L, Pfeiffer Jensen M, Hendricks O, et al. Self-protection strategies and health behaviour in patients with inflammatory rheumatic diseases during the COVID-19 pandemic: results and predictors in more than 12 000 patients with inflammatory rheumatic diseases followed in the Danish DANBIO registry. *RMD Open* 2021 Jan 05;7(1):e001505 [FREE Full text] [doi: [10.1136/rmdopen-2020-001505](https://doi.org/10.1136/rmdopen-2020-001505)] [Medline: [33402443](https://pubmed.ncbi.nlm.nih.gov/33402443/)]
24. Khader MA, Jabeen T, Namuju R. A cross sectional study reveals severe disruption in glycemic control in people with diabetes during and after lockdown in India. *Diabetes Metab Syndr* 2020 Nov;14(6):1579-1584 [FREE Full text] [doi: [10.1016/j.dsx.2020.08.011](https://doi.org/10.1016/j.dsx.2020.08.011)] [Medline: [32858476](https://pubmed.ncbi.nlm.nih.gov/32858476/)]
25. Miller WR, Von Gaudecker J, Tanner A, Buelow JM. Epilepsy self-management during a pandemic: Experiences of people with epilepsy. *Epilepsy Behav* 2020 Oct;111:107238 [FREE Full text] [doi: [10.1016/j.yebeh.2020.107238](https://doi.org/10.1016/j.yebeh.2020.107238)] [Medline: [32593874](https://pubmed.ncbi.nlm.nih.gov/32593874/)]

26. Nieto R, Pardo R, Sora B, Feliu-Soler A, Luciano JV. Impact of COVID-19 Lockdown Measures on Spanish People with Chronic Pain: An Online Study Survey. *J Clin Med* 2020 Nov 05;9(11):3558 [FREE Full text] [doi: [10.3390/jcm9113558](https://doi.org/10.3390/jcm9113558)] [Medline: [33167322](https://pubmed.ncbi.nlm.nih.gov/33167322/)]
27. Pal R, Yadav U, Verma A, Bhadada SK. Awareness regarding COVID-19 and problems being faced by young adults with type 1 diabetes mellitus amid nationwide lockdown in India: A qualitative interview study. *Prim Care Diabetes* 2021 Feb;15(1):10-15 [FREE Full text] [doi: [10.1016/j.pcd.2020.07.001](https://doi.org/10.1016/j.pcd.2020.07.001)] [Medline: [32660907](https://pubmed.ncbi.nlm.nih.gov/32660907/)]
28. Philip K, Cumella A, Farrington-Douglas J, Laffan M, Hopkinson N. Respiratory patient experience of measures to reduce risk of COVID-19: findings from a descriptive cross-sectional UK wide survey. *BMJ Open* 2020 Sep 09;10(9):e040951 [FREE Full text] [doi: [10.1136/bmjopen-2020-040951](https://doi.org/10.1136/bmjopen-2020-040951)] [Medline: [32912958](https://pubmed.ncbi.nlm.nih.gov/32912958/)]
29. Singh K, Kondal D, Mohan S, Jaganathan S, Deepa M, Venkateshmurthy NS, et al. Health, psychosocial, and economic impacts of the COVID-19 pandemic on people with chronic conditions in India: a mixed methods study. *BMC Public Health* 2021 Apr 08;21(1):685 [FREE Full text] [doi: [10.1186/s12889-021-10708-w](https://doi.org/10.1186/s12889-021-10708-w)] [Medline: [33832478](https://pubmed.ncbi.nlm.nih.gov/33832478/)]
30. Tiwari A, Kumar D, Ansari MS, Chaubey SK, Gupta NR, Agarwal V, et al. Impact of lockdown on self-care management among patients with type 2 Diabetes Mellitus residing in Lucknow city, India – A cross-sectional study. *Clinical Epidemiology and Global Health* 2021 Apr;10:100703. [doi: [10.1016/j.cegh.2021.100703](https://doi.org/10.1016/j.cegh.2021.100703)]
31. Thorpe J, Ashby S, Hallab A, Ding D, Andraus M, Dugan P, COVID-19Epilepsy (COV-E) Study Group. Evaluating risk to people with epilepsy during the COVID-19 pandemic: Preliminary findings from the COV-E study. *Epilepsy Behav* 2021 Feb;115:107658 [FREE Full text] [doi: [10.1016/j.yebeh.2020.107658](https://doi.org/10.1016/j.yebeh.2020.107658)] [Medline: [33341393](https://pubmed.ncbi.nlm.nih.gov/33341393/)]
32. Ziadé N, el Kibbi L, Hmamouchi I, Abdulateef N, Halabi H, Hamdi W, et al. Impact of the COVID - 19 pandemic on patients with chronic rheumatic diseases: A study in 15 Arab countries. *Int J Rheum Dis* 2020 Sep 05;23(11):1550-1557. [doi: [10.1111/1756-185x.13960](https://doi.org/10.1111/1756-185x.13960)]
33. Bala R, Srivastava A, Potsangbam T, Anal L, Ningthoujam GD. Self care practices and psychological distress among diabetic patients in Manipur during COVID-19: A scenario from the North East. *Diabetes Metab Syndr* 2021 Jan;15(1):93-98 [FREE Full text] [doi: [10.1016/j.dsx.2020.12.015](https://doi.org/10.1016/j.dsx.2020.12.015)] [Medline: [33316765](https://pubmed.ncbi.nlm.nih.gov/33316765/)]
34. da Silva JM, Verlengia R, de Oliveira JJ, Ribeiro AGSV, Barbosa CGR, Stotzer US, et al. Associations between sociodemographic factors and physical activity and sedentary behaviors in adults with chronic diseases during COVID-19 pandemic. *Sports Med Health Sci* 2020 Dec;2(4):216-220 [FREE Full text] [doi: [10.1016/j.smhs.2020.11.002](https://doi.org/10.1016/j.smhs.2020.11.002)] [Medline: [34189486](https://pubmed.ncbi.nlm.nih.gov/34189486/)]
35. Elran-Barak R, Mozeikov M. One Month into the Reinforcement of Social Distancing due to the COVID-19 Outbreak: Subjective Health, Health Behaviors, and Loneliness among People with Chronic Medical Conditions. *Int J Environ Res Public Health* 2020 Jul 27;17(15):5403 [FREE Full text] [doi: [10.3390/ijerph17155403](https://doi.org/10.3390/ijerph17155403)] [Medline: [32727103](https://pubmed.ncbi.nlm.nih.gov/32727103/)]
36. Flint SW, Brown A, Tahrani AA, Piotrkowicz A, Joseph A. Cross-sectional analysis to explore the awareness, attitudes and actions of UK adults at high risk of severe illness from COVID-19. *BMJ Open* 2020 Dec 29;10(12):e045309 [FREE Full text] [doi: [10.1136/bmjopen-2020-045309](https://doi.org/10.1136/bmjopen-2020-045309)] [Medline: [33376185](https://pubmed.ncbi.nlm.nih.gov/33376185/)]
37. Grabia M, Markiewicz-Żukowska R, Puścion-Jakubik A, Bielecka J, Nowakowski P, Gromkowska-Kępa K, et al. The Nutritional and Health Effects of the COVID-19 Pandemic on Patients with Diabetes Mellitus. *Nutrients* 2020 Sep 30;12(10):3013 [FREE Full text] [doi: [10.3390/nu12103013](https://doi.org/10.3390/nu12103013)] [Medline: [33008059](https://pubmed.ncbi.nlm.nih.gov/33008059/)]
38. Kovil R, Shah T, Chawla M, Karkhanis S, Padhye D, Sanghvi A, et al. Patient reported changes in metabolic health during COVID-19 induced lockdown – a cross sectional digital connect survey in type 2 diabetes. *EJEA* 2020 Aug 21. [doi: [10.1530/endoabs.70.EP554](https://doi.org/10.1530/endoabs.70.EP554)]
39. López-Sánchez GF, López-Bueno R, Gil-Salmerón A, Zauder R, Skalska M, Jastrzębska J, et al. Comparison of physical activity levels in Spanish adults with chronic conditions before and during COVID-19 quarantine. *Eur J Public Health* 2021 Feb 01;31(1):161-166 [FREE Full text] [doi: [10.1093/eurpub/ckaa159](https://doi.org/10.1093/eurpub/ckaa159)] [Medline: [32761181](https://pubmed.ncbi.nlm.nih.gov/32761181/)]
40. Mun C, Campbell C, McGill L, Aaron R. The Early Impact of COVID-19 on Chronic Pain: A Cross-Sectional Investigation of a Large Online Sample of Individuals with Chronic Pain in the United States, April to May, 2020. *Pain Med* 2021 Feb 23;22(2):470-480 [FREE Full text] [doi: [10.1093/pm/pnaa446](https://doi.org/10.1093/pm/pnaa446)] [Medline: [33537764](https://pubmed.ncbi.nlm.nih.gov/33537764/)]
41. Nachimuthu S, Vijayalakshmi R, Sudha M, Viswanathan V. Coping with diabetes during the COVID - 19 lockdown in India: Results of an online pilot survey. *Diabetes Metab Syndr* 2020 Jul;14(4):579-582 [FREE Full text] [doi: [10.1016/j.dsx.2020.04.053](https://doi.org/10.1016/j.dsx.2020.04.053)] [Medline: [32416527](https://pubmed.ncbi.nlm.nih.gov/32416527/)]
42. Pleguezuelos E, Del Carmen A, Moreno E, Ortega P, Vila X, Ovejero L, et al. The Experience of COPD Patients in Lockdown Due to the COVID-19 Pandemic. *COPD* 2020 Oct;Volume 15:2621-2627. [doi: [10.2147/copd.s268421](https://doi.org/10.2147/copd.s268421)]
43. Rogers NT, Waterlow NR, Brindle H, Enria L, Eggo RM, Lees S, et al. Behavioral Change Towards Reduced Intensity Physical Activity Is Disproportionately Prevalent Among Adults With Serious Health Issues or Self-Perception of High Risk During the UK COVID-19 Lockdown. *Front Public Health* 2020 Sep 30;8:575091 [FREE Full text] [doi: [10.3389/fpubh.2020.575091](https://doi.org/10.3389/fpubh.2020.575091)] [Medline: [33102424](https://pubmed.ncbi.nlm.nih.gov/33102424/)]
44. Saqib MAN, Siddiqui S, Qasim M, Jamil MA, Rafique I, Awan UA, et al. Effect of COVID-19 lockdown on patients with chronic diseases. *Diabetes Metab Syndr* 2020 Nov;14(6):1621-1623 [FREE Full text] [doi: [10.1016/j.dsx.2020.08.028](https://doi.org/10.1016/j.dsx.2020.08.028)] [Medline: [32889403](https://pubmed.ncbi.nlm.nih.gov/32889403/)]

45. Sauchelli S, Bradley J, England C, Searle A, Whitmarsh A. Exploring support needs of people living with diabetes during the coronavirus COVID-19 pandemic: insights from a UK survey. *BMJ Open Diabetes Res Care* 2021 Jun;9(1):1 [FREE Full text] [doi: [10.1136/bmjdr-2021-002162](https://doi.org/10.1136/bmjdr-2021-002162)] [Medline: [34099440](https://pubmed.ncbi.nlm.nih.gov/34099440/)]
46. Olickal JJ, Chinnakali P, Suryanarayana B, Ulaganeethi R, Kumar SS, Saya GK. Effect of COVID19 pandemic and national lockdown on persons with diabetes from rural areas availing care in a tertiary care center, southern India. *Diabetes Metab Syndr* 2020 Nov;14(6):1967-1972 [FREE Full text] [doi: [10.1016/j.dsx.2020.10.010](https://doi.org/10.1016/j.dsx.2020.10.010)] [Medline: [33059300](https://pubmed.ncbi.nlm.nih.gov/33059300/)]
47. Batista S, Souza A, Nogueira J, Andrade F, Thumé E, Teixeira D, et al. Protective behaviors for COVID-19 among Brazilian adults and elderly living with multimorbidity: the ELSI-COVID-19 initiative. *Cad Saude Publica* 2020;36Suppl 3(Suppl 3):e00196120 [FREE Full text] [doi: [10.1590/0102-311X00196120](https://doi.org/10.1590/0102-311X00196120)] [Medline: [33206836](https://pubmed.ncbi.nlm.nih.gov/33206836/)]
48. Galica J, Liu Z, Kain D, Merchant S, Booth C, Koven R, et al. Coping during COVID-19: a mixed methods study of older cancer survivors. *Support Care Cancer* 2021 Jun 06;29(6):3389-3398 [FREE Full text] [doi: [10.1007/s00520-020-05929-5](https://doi.org/10.1007/s00520-020-05929-5)] [Medline: [33404813](https://pubmed.ncbi.nlm.nih.gov/33404813/)]
49. Javanparast S, Roeger L, Kwok Y, Reed RL. The experience of Australian general practice patients at high risk of poor health outcomes with telehealth during the COVID-19 pandemic: a qualitative study. *BMC Fam Pract* 2021 Apr 08;22(1):69 [FREE Full text] [doi: [10.1186/s12875-021-01408-w](https://doi.org/10.1186/s12875-021-01408-w)] [Medline: [33832422](https://pubmed.ncbi.nlm.nih.gov/33832422/)]
50. Pati S, Mahapatra P, Kanungo S, Uddin A, Sahoo KC. Managing Multimorbidity (Multiple Chronic Diseases) Amid COVID-19 Pandemic: A Community Based Study From Odisha, India. *Front Public Health* 2020 Feb 1;8:584408 [FREE Full text] [doi: [10.3389/fpubh.2020.584408](https://doi.org/10.3389/fpubh.2020.584408)] [Medline: [33598442](https://pubmed.ncbi.nlm.nih.gov/33598442/)]
51. Wong SYS, Zhang D, Sit RWS, Yip BHK, Chung RY, Wong CKM, et al. Impact of COVID-19 on loneliness, mental health, and health service utilisation: a prospective cohort study of older adults with multimorbidity in primary care. *Br J Gen Pract* 2020 Sep 28;70(700):e817-e824. [doi: [10.3399/bjgp20x713021](https://doi.org/10.3399/bjgp20x713021)]
52. Horrell LN, Hayes S, Herbert LB, MacTurk K, Lawhon L, Valle CG, et al. Telemedicine Use and Health-Related Concerns of Patients With Chronic Conditions During COVID-19: Survey of Members of Online Health Communities. *J Med Internet Res* 2021 Feb 18;23(2):e23795 [FREE Full text] [doi: [10.2196/23795](https://doi.org/10.2196/23795)] [Medline: [33539307](https://pubmed.ncbi.nlm.nih.gov/33539307/)]
53. Burton BS, Patterson J, Robinson M, Patel D, Allison M, Brassil KJ. Insights from individuals with chronic conditions in the context of COVID-19. *Patient Experience Journal* 2020 Aug 04;7(2):39-52. [doi: [10.35680/2372-0247.1489](https://doi.org/10.35680/2372-0247.1489)]
54. Clemente-Suárez VJ, Ramos-Campo DJ, Mielgo-Ayuso J, Dalamitros AA, Nikolaidis PA, Hormeño-Holgado A, et al. Nutrition in the Actual COVID-19 Pandemic. A Narrative Review. *Nutrients* 2021 Jun 03;13(6):1924 [FREE Full text] [doi: [10.3390/nu13061924](https://doi.org/10.3390/nu13061924)] [Medline: [34205138](https://pubmed.ncbi.nlm.nih.gov/34205138/)]
55. Tisminetzky M, Delude C, Hebert T, Carr C, Goldberg R, Gurwitz J. Age, Multiple Chronic Conditions, and COVID-19: A Literature Review. *J Gerontol A Biol Sci Med Sci* 2022 Apr 01;77(4):872-878 [FREE Full text] [doi: [10.1093/gerona/glaa320](https://doi.org/10.1093/gerona/glaa320)] [Medline: [33367606](https://pubmed.ncbi.nlm.nih.gov/33367606/)]
56. Haybar H, Kazemnia K, Rahim F. Underlying Chronic Disease and COVID-19 Infection: A State-of-the-Art Review. *Jundishapur J Chronic Dis Care* 2020 Apr 15;9(2). [doi: [10.5812/jjcdc.103452](https://doi.org/10.5812/jjcdc.103452)]
57. Figliozzi S, Masci PG, Ahmadi N, Tondi L, Koutli E, Aimo A, et al. Predictors of adverse prognosis in COVID-19: A systematic review and meta-analysis. *Eur J Clin Invest* 2020 Oct 27;50(10):e13362. [doi: [10.1111/eci.13362](https://doi.org/10.1111/eci.13362)] [Medline: [32726868](https://pubmed.ncbi.nlm.nih.gov/32726868/)]

Abbreviations

PA: physical activities

PRISMA: Preferred Reporting Items for Systematic Reviews and Meta-Analyses

PROSPERO: International Prospective Register of Systematic Reviews

Edited by A Kushniruk; submitted 27.11.21; peer-reviewed by M Herron, J Parsons; comments to author 10.01.22; revised version received 18.02.22; accepted 22.04.22; published 21.06.22.

Please cite as:

Sakur F, Ward K, Khatri NN, Lau AYS

Self-care Behaviors and Technology Used During COVID-19: Systematic Review

JMIR Hum Factors 2022;9(2):e35173

URL: <https://humanfactors.jmir.org/2022/2/e35173>

doi: [10.2196/35173](https://doi.org/10.2196/35173)

PMID: [35442904](https://pubmed.ncbi.nlm.nih.gov/35442904/)

©Fareeya Sakur, Kanasha Ward, Neha Nafees Khatri, Annie Y S Lau. Originally published in *JMIR Human Factors* (<https://humanfactors.jmir.org>), 21.06.2022. This is an open-access article distributed under the terms of the Creative Commons

Attribution License (<https://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Human Factors, is properly cited. The complete bibliographic information, a link to the original publication on <https://humanfactors.jmir.org>, as well as this copyright and license information must be included.

Original Paper

Health Professionals' eHealth Literacy and System Experience Before and 3 Months After the Implementation of an Electronic Health Record System: Longitudinal Study

Lars Kayser¹, MD, PhD; Astrid Karnoe¹, PhD; Emily Duminski¹, BSc; Svend Jakobsen², MSSc; Rikke Terp², RN, PhD; Susanne Dansholm², BSN, RN; Michael Roeder², MD, DMSc; Gustav From², MD, PhD

¹Section of Health Service Research, Department of Public Health, University of Copenhagen, Copenhagen, Denmark

²Department of Internal Medicine, Herlev-Gentofte Hospital, Hellerup, Denmark

Corresponding Author:

Lars Kayser, MD, PhD

Section of Health Service Research

Department of Public Health

University of Copenhagen

Øster Farimagsgade 5

Copenhagen, 1353

Denmark

Phone: 45 28757291

Email: lk@sund.ku.dk

Abstract

Background: The implementation of an integrated electronic health record (EHR) system can potentially provide health care providers with support standardization of patient care, pathways, and workflows, as well as provide medical staff with decision support, easier access, and the same interface across features and subsystems. These potentials require an implementation process in which the expectations of the medical staff and the provider of the new system are aligned with respect to the medical staff's knowledge and skills, as well as the interface and performance of the system. Awareness of the medical staff's level of eHealth literacy may be a way of understanding and aligning these expectations and following the progression of the implementation process.

Objective: The objective of this study was to investigate how a newly developed and modified instrument measuring the medical staff's eHealth literacy (staff eHealth Literacy Questionnaire [eHLQ]) can be used to inform the system provider and the health care organization in the implementation process and evaluate whether the medical staff's perceptions of the ease of use change and how this may be related to their level of eHealth literacy.

Methods: A modified version of the eHLQ was distributed to the staff of a medical department in Denmark before and 3 months after the implementation of a new EHR system. The survey also included questions related to users' perceived ease of use and their self-reported information technology skills.

Results: The mean age of the 194 participants before implementation was 43.1 (SD 12.4) years, and for the 198 participants after implementation, it was 42.3 (SD 12.5) years. After the implementation, the only difference compared with the preimplementation data was a small decrease in staff eHLQ5 (*motivated to engage with digital services*; unpaired 2-tailed *t* test; $P=.009$; effect size 0.267), and the values of the scales relating to the medical staff's knowledge and skills (eHLQ1-3) were approximately ≥ 3 both before and after implementation. The range of scores was narrower after implementation, indicating that some of those with the lowest ability benefited from the training and new experiences with the EHR. There was an association between perceived ease of use and the 3 tested staff eHLQ scales, both before and after implementation.

Conclusions: The staff eHLQ may be a good candidate for monitoring the medical staff's digital competence in and response to the implementation of new digital solutions. This may enable those responsible for the implementation to tailor efforts to the specific needs of segments of users and inform them if the process is not going according to plan with respect to the staff's information technology-related knowledge and skills, trust in data security, motivation, and experience of a coherent system that suits their needs and supports the workflows and data availability.

(*JMIR Hum Factors* 2022;9(2):e29780) doi:[10.2196/29780](https://doi.org/10.2196/29780)

KEYWORDS

health care professionals; eHealth literacy; electronic health record; implementation; digital health; eHealth; health literacy; health records; eHealth records; patient care

Introduction

Background

During the past 50 years, technological and digital evolution has facilitated the transformation of the organization and delivery of health services [1]. For more than two decades, it has been anticipated that electronic health records (EHRs), also termed electronic medical records or electronic patient records, would provide more efficient, effective, and safe workflows, benefiting both providers and patients [2-8].

In many regions, primarily Europe and the United States, the first generations of EHR were launched in the late '90s or the early 2000s [8,9]. These systems have been either upgraded or replaced with new systems with more advanced features and the integration of functions from various specialized systems into one system. The new generations of EHR have the potential to support standardization of patient care, pathways, and workflows, as well as provide organizations with data for business intelligence and health care professionals with decision support, easier access, and the same interface across its features and subsystems [8].

Factors for Medical Staff's Acceptance of an EHR

Not all implementations of EHR systems have been successful over the years. The reasons for this may be understood in relation to the context, content, and processes of EHR [9]. This includes the structure and digital maturity of the organizations, influence on or interaction with existing workflows, degree of involvement, digital experience, and competence of the staff [2,4,8-10].

The attitude of the medical staff toward a new EHR, as well as their capability to gain benefits, is related to the staff's level of information technology (IT) or eHealth literacy (eHL) [4,8,10]. To accept and adopt the technology, the user needs to feel confident and expect a good user experience based on the perceived ease of use and usefulness [4].

In general, most studies on the adoption and acceptance of technology build on the Technology Acceptance Model (TAM) or the Unified Theory of Acceptance and Use of Technology (UTAUT) [4,11,12]. Both TAM and UTAUT are relatively old theories that have not been specifically developed for the health care sector but have been adapted in several studies in the context of health [4]; for example, performance expectancy (individuals believe that the use of technology will be beneficial), effort expectancy (expected ease of use), social influence (expected attitude of significant others toward using the technology), and facilitating conditions (organizational or technical resources and preconditions to technology use) [10].

In 2015, Monkman and Kushniruk [13] proposed the *Consumer Health Information System Adoption Model*. The model is based on a theoretical approach and suggests, in alignment with the TAM and UTAUT, that an essential factor for adoption is the

user experience; more importantly, they proposed that user experience relates not only to usability and perceived ease of use or usefulness but also to the individual user's level of eHL as the user's level affects their perceived user experience and influences the requirements of the systems interface [13].

eHealth Literacy

eHL was introduced to describe the competences needed to engage with digital health solutions in a health context. eHL was originally conceptualized by Norman and Skinner [14], who also proposed the first definition: "the ability to seek, find, understand, and appraise health information from electronic sources and apply the gained knowledge to addressing or solving a health problem" [14]. In 2015, Nørsgaard et al [15] challenged the original concept with the proposal of a new and more comprehensive model, the eHL Framework (eHLF). The eHLF comprises 7 dimensions that not only address the user's knowledge and skills, similar to Norman and Skinner [14], but also address the interface and context (ie, the user's trust, motivation, and experience with digital services and technology).

The dimensions relating to the user's knowledge and skills are eHLF1, *ability to process information*; eHLF2, *engagement in own health*; and eHLF3, *ability to actively engage with digital services*. The dimensions relating to the user's trust in the way their health data are handled and the benefits of digital services are eHLF4, *feel safe and in control*, and eHLF5, *motivated to engage with digital services*. The final two dimensions, eHLF6, *access to digital services that work*, and eHLF7, *digital services that suit individual needs*, relate to the experience of the available digital services in relation to access to relevant information whenever it is needed in a way that suits the user's needs [15]. The user's self-reported capability within the 7 dimensions can be quantified using the eHL Questionnaire (eHLQ), which is based on the eHLF [16].

Both the eHLF [17,18] and the eHLQ may, alone [19,20] or in combination with other scales such as the Readiness and Enablement Index for Health Technology instrument [21-23], help identify potential barriers or facilitators with respect to the user's capabilities, their trust and motivation, and their experiences with digital services.

In the context of the implementation of an EHR, assessment of eHL among medical staff has the potential to provide the supplier and health care organization before the implementation with insights into which groups may have particular needs to be addressed and after the implementation with insights into how the implementation, including educational programs, affects users' knowledge, skills, motivation, and experience.

The Setting of the Study

Part of the validation of the eHLQ was the inclusion of data from a medical outpatient clinic in the Capital Region of Denmark, Herlev-Gentofte University Hospital, from November 2015 to March 2016 [19]. Incidentally, one of the largest

implementations of an EHR in northern Europe was planned to take place simultaneously at the same hospital.

The new integrated EHR system was planned to replace >20 existing systems and be followed by investment in new technologies such as handheld devices, mobile computers, and standardized equipment such as infusion pumps. At the time of implementation, the medical staff was used to using a traditional EHR supporting documentation in notes along with a laboratory system, an imaging system, and a medicine prescription system, as well as a patient administrative system primarily used by medical secretaries and a documentation system for nurses. The latter was used by medical physicians, nursing assistants, and registered nurses. Everyday use was supported by both local health professionals trained as superusers and by regional IT support with a help desk.

The Capital Region of Denmark's expected outcomes of the introduction of the new integrated EHR were more efficient workflows that were better supported by technology and a reorientation of the professional roles and tasks, including easier and better communication with outpatients [24]. An important change in workflow was the introduction of the principle that

the individual staff member responsible for an order should also enter this into the system, which changed the work balance among medical physicians, registered nurses, and medical secretaries [24]. The introduction of an anticipated, easier to use EHR better supporting communication and workflows, together with a 3 full days training program for the nursing assistants and ≥ 4 full days for the other groups of medical staff, followed by 2 weeks of intensive support by superusers and specially trained *floor walkers* after the launch of the new EHR system, led us to expect that the overall effect of the implementation would be an increase in the medical staff's eHL and their perceived ease of use.

The combination of having a new instrument to assess the multifaceted dimensions of eHL during the implementation of a promising new suite of EHR components, as well as curiosity about how this would influence the medical staff's eHL profile, led us to initiate this study. Our aim is to evaluate the eHL of the medical staff using the eHLQ before and 3 months after the implementation of the EHR to examine the overall effect of the introduction of the new system.

We worked from the hypotheses presented in [Textbox 1](#).

Textbox 1. Hypotheses of this study.

Hypothesis 1

- Hypothesis 1.1: The medical staff's personal knowledge and skills (eHealth Literacy Questionnaire [eHLQ] 1-3) will increase as a consequence of the introduction of the new electronic health record (EHR) with a 3- to 4-day training program and extensive support for the first 2 weeks after implementation.
- Hypothesis 1.2: An overall positive experience with the new system with an EHR will improve the sense of feeling safe and in control (staff eHLQ4) and increase motivation (staff eHLQ5) as the medical staff experience the expected benefits of an integrated EHR system.
- Hypothesis 1.3: The implementation will provide an experience of an EHR that brings data together, makes them easier to access (staff eHLQ6), and better suits the individual needs (staff eHLQ7).
- Hypothesis 1.4: The eHLQ scores may differ between the groups of medical staff because of different professional cultures, tasks, obligations, and responsibilities.

Hypothesis 2

- Hypothesis 2.1: The experience of ease of access, ease of data sharing, and stability of the information technology system will increase with the new integrated system running on a more stable platform.
- Hypothesis 2.2: The increase may be associated with staff eHLQ5, staff eHLQ6, and staff eHLQ7, establishing a possible association between factors known to be important for technology acceptance and eHealth literacy dimensions.

To explore these hypotheses, we formulated the following research questions (RQs):

- RQ1: What is the level of the medical staff's eHL before and 3 months after the implementation of the new EHR?
- RQ2: How do medical staff perceive ease of use, as measured by the ease of access, ease of data sharing, and stability of the existing EHR before implementation, compared with the new integrated EHR system after implementation, and are there any differences between professions?
- RQ3: Is there an association between the scores of staff eHLQ5-7 and perceived ease of use, as measured by ease of access, ease of data sharing, and stability of the system?

Methods

Overview

The study was originally designed as a longitudinal study to evaluate the medical staff's eHL, perceived ease of use, and use of functions before implementation and at 3 and 12 months after implementation. The involved department was restructured before month 12 by fusing with 2 other medical departments, resulting in a change of jobs for 3 of the 4 clinical working authors of this study and relocation of the acute clinical unit and other specialties such as gastroenterology from one location in the city of Gentofte to another location in the city of Herlev. Therefore, we had to exclude the 12-month follow-up, as it was not feasible for us to conduct. A planned complementary qualitative study was also not feasible in the initial period

because of a lack of support from a higher level of the organization responsible for the implementation.

The study was designed with 2 cross-sectional samples, inviting all the medical staff employed at 2 time points. In March 2016, an invitation was sent by email to the medical staff working in all units, including the outpatient clinic, at the Department of Medicine C, Herlev-Gentofte University Hospital, containing a link to the survey, and by mid-March, a reminder to participate was sent to those who did not initially respond. The second

survey was sent out in September 2016. The study was endorsed by the head of the department, who took an active part in recruiting respondents at both time points.

The survey was sent to 295 medical staff members in both rounds, with a response rate of 65.8% (194/295) in the first round and 67.1% (198/295) in the second round and respondents answering some or all questions. The distribution of respondents among different groups of medical staff is presented in [Table 1](#).

Table 1. Distribution of medical staff^a.

Staff	Before implementation (N=194), n (%)	After implementation (N=198), n (%)
Medical physician	46 (23.7)	50 (25.3)
Medical secretary	29 (14.9)	26 (13.1)
Nursing assistant	16 (8.2)	15 (7.6)
Registered nurse	97 (50)	104 (52.5)
Other professions	6 (3.1)	3 (1.5)

^aThe table includes respondents who answered some or all questions.

All groups of medical staff employed at the department were represented in response to the survey. For this study, we report on all respondents in relation to overall statistics but have not included the group of other professions (9/295, 3.1%) when reporting on groups of professional medical staff (ie, medical physicians, medical secretaries, nursing assistants, and registered nurses). In Denmark, these 4 professional groups have the qualifications and level of education according to the International Standard Classification of Education (ISCED) given in the following sections [25].

Medical physicians had a master's degree in medicine at ISCED level 7 [25]. Some of them also held a PhD or medical thesis degree at ISCED level 8. Their experience ranged from registrars leaving the medical school to consultants, who were specialists. Registered nurses had 3.5 years of education and held a bachelor's degree in nursing at ISCED level 6. Nursing assistants had a vocational education, which currently is 3 years at ISCED level 4; however, some of the respondents may have had a previous education of 2 years at ISCED level 3. Medical secretaries also had a vocational education of 3 years, with specialization in the medical field.

The surveys in the project were intentionally designed so that they would not be misinterpreted as an evaluation of the new EHR. In accordance with the hypotheses and RQs, the sole intention of the surveys was to describe the change in the medical staff's eHL and their perceived ease of use of the 2 different EHR solutions.

The survey comprised four sections: (1) digital experience; (2) the staff eHLQ; (3) experience of use with the EHR, including perceived ease of use; and (4) use of functions of the EHR. The use of functions and components will be reported elsewhere and are not included here.

Sex, age, and professional roles were extracted from the administrative system and merged with the survey. This was performed by an administrator based on each participant's

unique employee identifier. After the merging was complete, person-identifiable data were removed from the file, which was then handed over to the author group for analysis.

Digital Skills

As an indicator of experience with digital services in their private lives, the respondents were asked to report on their use of the national digital mail service called e-Boks. e-Boks facilitates all communication from public authorities in Denmark to citizens aged >15 years. Individuals with language difficulties or disabilities can be exempted from the e-Boks system. The respondents reported on their use with four response options—rarely or never, at least once every 6 months, at least once a month, and at least once a week—scored from 1 to 4.

The second question was how their colleagues would describe their user level in relation to the systems they used at work with three options—standard user, advanced user, or expert user—scored from 1 to 3.

Staff eHLQ

The staff eHLQ is a modified version of the eHLQ [16]. The modification comprised rephrasing 12 items in scales 4 to 7 of the eHLQ to change the perspective of the respondent from themselves to their interaction with patients; for example, item 24, which was changed from “I find I get better services from my healthcare provider when I use...” to “I find that patients receive better services from health professionals when...” The items in staff eHLQ1-3 are equivalent to the validated eHLQ, except that 1 item in eHLQ1 is missing because the staff eHLQ used here was based on an earlier version of the eHLQ.

Therefore, the staff eHLQ in this study comprised 34 items covering seven dimensions of eHL in the following scales: eHLQ1, *using technology to process health information*; eHLQ2, *understanding of health concepts and language*; eHLQ3, *ability to actively engage with digital services*; staff eHLQ4, *feeling safe and in control*; staff eHLQ5, *motivated to engage with digital services*; staff eHLQ6, *access to digital services that*

work, and staff eHLQ7, *digital services that suit individual needs*. The eHLQ1 and staff eHLQ7 scales comprise 4 items, eHLQ2 to staff eHLQ5 comprise 5 items, and staff eHLQ6 comprises 6 items. The response options ranged from strongly disagree to strongly agree and were scored from 1 to 4 [16].

The validation of the eHLQ was reported by Kayser et al [16]. To ensure that the aforementioned minor changes did not change the internal consistency, we calculated Cronbach α , which is similar to those initially reported with the following values: eHLQ1=.7519, eHLQ2=.7646, eHLQ3=.8413, eHLQ4=.7463, eHLQ5=.7422, eHLQ6=.6786, and eHLQ7=.8131.

Perceived Ease of Use Evaluated as the Experience of Use With the Digital Information and IT Systems

This part comprised three items adopted from a national, regular survey, *Termostaten*, administrated by *The Danish e-Observatory* [26], which assesses users' self-reported experience of the following three items:

1. Quick and easy access: "In my daily work I have quick and easy access to all the essential digital information from my own sector or unit (department or hospital) that I need."
2. Sharing of data to reduce double registration: "In my daily work I experience, that data is shared between systems in a way that makes double registrations avoidable."
3. Stability of systems: "In my daily work I experience, that the work-related IT-system I use every day are stable and function without operational problems or crashes."

The 3 items are all considered to report on perceived ease of use and are used for this purpose in the analysis. The response options ranged from strongly disagree to strongly agree, with scores ranging from 1 to 4. Each of the 3 items was evaluated separately.

Statistical Analysis

We treated the 2 samples as independent in the analysis as the questionnaire was administered anonymously to us.

To test hypotheses 1.1, 1.2, 1.3, and 2.1, we used an unpaired 2-tailed *t* test to compare the levels of eHL and perceived ease of use before and after the implementation. The effect size was calculated as Cohen *d*, and 1-way ANOVA was used to examine significant differences between the 4 medical staff groups in terms of their scores. The Tukey honest significant difference test was used a posteriori to determine which medical staff groups' means differed significantly from each other. We also used an unpaired *t* test to examine differences in scores between males and females. Pearson *r* was used to examine the association between age in relation to the eHLQ scales and self-reported IT skills.

To test hypothesis 2.2, Pearson *r* was calculated to examine the correlations among the experience of quick and easy access; sharing of data to reduce double registration; and the stability of the IT system; and staff eHLQ5, staff eHLQ6, and staff eHLQ7.

All quantitative data are reported as means and SDs.

Statistical calculations were performed using Stata (version 16; StataCorp).

Ethics Approval

Under Danish law, permission from an ethics committee was not required as biological material was not obtained or processed in the study, and no clinical intervention of the respondents was performed. The data were gathered by the hospital administration and stored on their servers. The anonymized data were further processed at the University of Copenhagen. Before data collection, all respondents were introduced to the survey by their local leaders. When initiating the survey, the respondents provided informed consent to participate by filling in the survey.

Results

Overview

The age and sex distributions of the 2 samples are presented in Table 2. The mean age of the sample before implementation was 43.1 (SD 12.4) years and 42.3 (SD 12.5) years in the sample after implementation. The sample mainly comprised female respondents.

Most respondents used the national email system, e-Boks, regularly. Few medical physicians and nursing assistants did not use the national email service before the implementation of the EHR. After implementation, all groups used the service at least once every 6 months, and most of them used the service more regularly. The average scores were approximately the same before and after implementation (Table 3). The score of how a colleague described their IT skills did not change over time and did not differ between the medical staff groups. Before implementation, there was a minor negative correlation with age ($r=-0.1965$; $P=.009$), which increased 3 months after implementation ($r=-0.283$; $P<.001$), signifying that the younger members of staff were more confident in their IT skills, a difference that increased after the introduction of the new EHR system. We also found a difference in males scoring higher than females both before implementation (mean 2.074, SD 0.675 vs mean 1.516, SD 0.661; $P<.001$; effect size 0.83) and 3 months after implementation (mean 1.848, SD 0.712 vs mean 1.538, SD 0.627; $P=.01$; effect size 0.482).

Table 2. Background variables by job functions.

Characteristics and staff	Before implementation (N=194)	After implementation (N=198)
Age (years), mean (SD; range)		
Overall	43.1 (12.4; 23-68)	42.3 (12.5; 24-68)
Medical physician	43.6 (12.7; 26-68)	42.0 (13.0; 27-68)
Medical secretary	49.3 (10.7; 24-64)	50.1 (9.8; 25-63)
Nursing assistant	52.4 (8.7; 34-64)	54.8 (9.0; 35-66)
Registered nurse	39.3 (12.0; 23-66)	38.8 (11.3; 30-60)
Sex, n (%)		
Overall		
Male	29 (14.9)	35 (17.7)
Female	165 (85.1)	163 (82.3)
Medical physician		
Male	16 (8.2)	22 (11.1)
Female	30 (15.5)	28 (14.1)
Medical secretary		
Male	2 (1)	0 (0)
Female	27 (13.9)	26 (13.1)
Nursing assistant		
Male	1 (0.5)	2 (1)
Female	15 (7.7)	13 (6.6)
Registered nurse		
Male	10 (5.2)	10 (5.1)
Female	87 (44.8)	94 (47.5)

Table 3. Information technology skills by job functions.

Skill and staff	Before implementation, mean (SD; range)	After implementation, mean (SD; range)
Use of e-Boks		
Medical physician	3.301 (0.615; 1-4)	3.354 (0.558; 2.0-4.0)
Medical secretary	3.217 (0.629; 2-4)	3.3 (0.544; 2.0-4.0)
Nursing assistant	3.310 (0.660; 1-4)	3.35 (0.485; 3.0-4.0)
Nursing assistant	3.375 (0.806; 1-4)	3.67 (0.488; 3.0-4.0)
Registered nurse	3.333 (0.556; 2-4)	3.346 (0.587; 2.0-4.0)
Information technology skills described by colleague		
Medical physician	1.601 (0.692; 1.0-3.0)	1.594 (0.65; 1.0-3.0)
Medical physician	1.667 (0.6396; 1.0-3.0)	1.553 (0.619; 1.0-3.0)
Medical secretary	1.963 (0.854; 1.0-3.0)	1.783 (0.736; 1.0-3.0)
Nursing assistant	1.2 (0.414; 1.0-2.0)	1.286 (0.726; 1.0-3.0)
Registered nurse	1.547 (0.663; 1.0-3.0)	1.61 (0.618; 1.0-3.0)

eHLQ Scales

After 3 months from the implementation, the only difference compared with the preimplementation data was a decrease in staff eHLQ5 (*motivated to engage with digital services*; unpaired *t* test; $P=.009$; effect size 0.267), whereas the other scales did not differ from before implementation (effect size ranging from 0.0093 to 0.0916). As seen in Table 4, the eHLQ scores in relation to the respondents' knowledge and skills (eHLQ1-3)

were approximately ≥ 3 both before and after implementation. The range of scores was narrower after the implementation, indicating that some of those with the lowest ability benefited from training and new experiences with the EHR. On the basis of these findings, we rejected hypotheses 1.1 to 1.3. With respect to hypothesis 1.4, we found differences among the groups of medical staff for some of the scales, both before and after the

implementation of the EHR, which partly supports our hypothesis.

Table 4. eHLQ^a scores by professional groups.

Scales and staff	Before implementation, mean (SD; range)	After implementation, mean (SD; range)
eHLQ1: using technology to process health information	2.980 (0.597; 1.0-4.0)	3.009 (0.574; 1.5-4.0)
Medical physician	2.989 (0.570; 1.75-4.0)	2.893 (0.663; 1.5-4.0)
Medical secretary	2.896 (0.611; 1.0-4.0)	2.860 (0.479; 2.0-4.0)
Nursing assistant	2.921 (0.778; 1.0-4.0)	3.233 (0.458; 2.3-4.0)
Registered nurse	3.018 (0.565; 1.5-4.0)	3.046 (0.544; 2.0-4.0)
eHLQ2: understanding of health concepts and language	3.399 (0.467; 1.0-4.0)	3.407 (0.439; 2.0-4.0)
Medical physician	3.565 (0.356; 2.6-4.0)	3.551 (0.429; 2.0-4.0)
Medical secretary	3.255 (0.487; 2.4-4.0)	3.160 (0.374; 2.4-3.8)
Nursing assistant	3.163 (0.742; 1.0-4.0)	3.413 (0.389; 3.0-4.0)
Registered nurse	3.408 (0.426; 2.4-4.0)	3.389 (0.441; 2.0-4.0)
eHLQ3: ability to actively engage with digital services	3.359 (0.505; 1.0-4.0)	3.364 (0.502; 1.8-4.0)
Medical physician	3.448 (0.458; 2.4-4.0)	3.473 (0.493; 2.4-4.0)
Medical secretary	3.407 (0.559; 2.4-4.0)	3.176 (0.601; 1.8-4.0)
Nursing assistant	3.188 (0.675; 1.0-4.0)	3.227 (0.345; 2.6-3.6)
Registered nurse	3.333 (0.466; 2.2-4.0)	3.363 (0.487; 2.0-4.0)
Staff eHLQ4: feel safe and in control	2.953 (0.418; 1.8-4.0)	2.914 (0.418; 1.0-4.0)
Medical physician	2.843 (0.436; 1.8-4.0)	2.838 (0.491; 1.0-3.8)
Medical secretary	3.069 (0.461; 2.2-4.0)	2.912 (0.451; 2.0-3.8)
Nursing assistant	2.987 (0.325; 2.2-3.6)	2.960 (0.275; 2.6-3.8)
Registered nurse	2.962 (0.407; 1.8-4.0)	2.934 (0.379; 2.0-4.0)
Staff eHLQ5: motivated to engage with digital	2.783 (0.445; 1.6-3.8)	2.665 (0.439; 1.4-4.0)
Medical physician	2.839 (0.482; 1.6-3.8)	2.675 (0.486; 1.4-4.0)
Medical secretary	2.821 (0.379; 2.0-3.6)	2.696 (0.487; 2.0-3.6)
Nursing assistant	2.880 (0.477; 1.8-3.8)	2.880 (0.413; 2.0-3.8)
Registered nurse	2.738 (0.446; 1.6-3.8)	2.604 (0.395; 1.6-4.0)
Staff eHLQ6: access to digital services that work	2.566 (0.403; 1.5-3.8)	2.603 (0.411; 1.3-4.0)
Medical physician	2.391 (0.461; 1.5-3.5)	2.417 (0.427; 1.3-3.3)
Medical secretary	2.661 (0.338; 2.2-3.7)	2.607 (0.333; 2.2-3.7)
Nursing assistant	2.778 (0.325; 2.0-3.3)	2.833 (0.383; 2.2-3.7)
Registered nurse	2.589 (0.381; 1.7-3.8)	2.642 (0.393; 1.3-4.0)
Staff eHLQ7: digital services that suit individual needs	2.549 (0.508; 1.0-4.0)	2.510 (0.506; 1.3-4.0)
Medical physician	2.321 (0.499; 1.0-3.5)	2.229 (0.489; 1.3-3.0)
Medical secretary	2.741 (0.381; 2.0-3.5)	2.470 (0.435; 2.0-3.3)
Nursing assistant	2.783 (0.352; 1.8-3.0)	2.800 (0.368; 2.0-3.3)
Registered nurse	2.572 (0.535; 1.0-4.0)	2.597 (0.498; 1.3-4.0)

^aeHLQ: eHealth Literacy Questionnaire.

Before Implementation

Before implementation, the score of eHLQ2 (*understanding of health concepts and language*) showed significant differences among the 4 groups ($F_{3,185}=4.47$; $P=.005$). The medical

physicians scored significantly higher than the medical secretaries (Tukey test, $P=.02$) and nursing assistants (Tukey test, $P=.01$). There were no significant differences among the groups for eHLQ1 (*using technology to process health*

information) and eHLQ3 (ability to actively engage with digital services).

The number of respondents who scored lower than two-thirds of the maximum score (2.67) varied between 22% (10/46) and 31% (5/16) among the staff groups for eHLQ1 (using technology to process health information), with medical physicians representing the lowest and nursing assistants the highest percentage. For eHLQ2 (understanding of health concepts and language), the percentage varied from 2% (1/46) to 17% (5/29) and in eHLQ3 (the ability to actively engage with digital services), from 6% (1/16) to 10% (3/29), with medical physicians representing the lowest percentage again but now with the medical secretaries representing the highest percentage <2.67 in both scales.

The scores of staff eHLQ4 (feel safe and in control) and staff eHLQ5 (motivated to engage with digital services), which relate to the perception of the use of the system, were lower than the scores in eHLQ1 to eHLQ3. There were no differences between the groups.

With regards to staff eHLQ6 (access to digital services that work) and staff eHLQ7 (digital services that suit individual needs), which both reflect an overall experience with digital health services, the scores were even lower. Before implementation, for staff eHLQ6 (access to digital services that work), there were significant differences between the groups ($F_{3,183}=5.16$; $P=.002$). Medical physicians had a significantly lower score than medical secretaries (Tukey test, $P=.02$), nursing assistants (Tukey test, $P=.006$), and registered nurses (Tukey test, $P=.03$). The abovementioned findings do not appear to be associated with differences in age or sex among the groups, as the only association between age and eHLQ scores was a small negative correlation for eHLQ3 (ability to actively engage with digital services; $r=-0.2158$; $P=.003$) and between males and females for staff eHLQ5 (motivated to engage with digital services; mean 3.027, SD 0.477 vs mean 2.739, SD 0.426; $P=.001$; effect size 0.637).

After Implementation

After implementation, the scores of eHLQ2 (understanding of health concepts and language) differed among the professional groups ($F_{3,191}=4.72$; $P=.003$), where medical physicians had a significantly higher score than medical secretaries (Tukey test, $P=.001$).

In addition, for staff eHLQ6 (access to digital services that work), the ANOVA test showed significant differences among the groups ($F_{3,190}=5.61$; $P=.001$), where the medical physicians had a lower score than the nursing assistants (Tukey test, $P=.002$) and the registered nurses (Tukey test, $P=.007$). This pattern was repeated for the postimplementation measurement of staff eHLQ7 (digital services that suit individual needs), with significant differences among the groups ($F_{3,190}=8.51$; $P<.001$), where the medical physicians had a significantly lower score than the nursing assistants (Tukey test, $P<.001$) and the registered nurses (Tukey test, $P<.001$).

After implementation, there was a negative correlation between age and four of the seven eHLQ scales: eHLQ1, using technology to process health information ($r=-0.193$; $P=.007$); eHLQ2, understanding of health concepts and language ($r=-0.147$; $P=.04$); eHLQ3, ability to actively engage with digital services ($r=-0.263$; $P<.001$); and staff eHLQ4, feel safe and in control ($r=-0.153$; $P=.04$). There was a difference between males and females in the eHLQ2 (understanding of health concepts and language; mean 3.548, SD 0.527 vs mean 3.376, SD 0.413; $P=.04$; effect size 0.394) and staff eHLQ5 (motivated to engage with digital services; mean 2.834, SD 0.500 vs mean 2.627, SD 0.417; $P=.01$; effect size 0.478).

Association Between Self-reported Skills and eHLQ

To support the content validity, we tested whether there were any associations between eHLQ1 to eHLQ3 and what the respondent believed a colleague would describe their IT skills by calculating Pearson r . For the measurements before implementation, there were moderate to strong correlations among the three eHLQ scales and the IT skills item: eHLQ1 (using technology to process health information; $r=0.2176$; $P=.004$), eHLQ2 (understanding of health concepts and language; $r=0.2522$; $P<.001$), and eHLQ3 (ability to actively engage with digital services; $r=0.4471$; $P<.001$).

For the postimplementation measurements, there were similar correlations among the three eHLQ scales and the IT skills item: eHLQ1 (using technology to process health information; $r=0.1926$; $P=.008$), eHLQ2 (understanding of health concepts and language; $r=0.2244$; $P=.002$), and eHLQ3 (ability to actively engage with digital services; $r=0.4429$; $P<.001$).

This may be associated with the age of the respondents, as we also found a negative correlation between age and these 3 scales, as well as for the IT skills scale, as reported previously.

Perceived Ease of Use

The respondents scored the lowest on the stability of IT system items before implementation (Table 5). This item was the only one to increase after implementation of the new EHR system, whereas the 2 others did not change, which partly confirms hypothesis 2.1, that the new EHR system would increase the perceived ease of use. When comparing the groups of medical staff before implementation for the item regarding sharing of data to reduce doublet registration, there were significant differences among the groups ($F_{3,185}=5.24$; $P=.002$). Here, medical physicians had a lower score than medical secretaries (Tukey test, $P=.005$) and nursing assistants (Tukey test, $P=.04$). In addition, the registered nurses had a lower score than the medical secretaries (Tukey test, $P=.04$).

When comparing the groups of medical staff after implementation, there was still an overall significant difference between the groups with respect to their experience of data being shared between IT systems to reduce doublet registration ($F_{3,191}=7.48$; $P<.001$). Medical physicians had a significantly lower score than nursing assistants (Tukey test, $P<.001$) and registered nurses (Tukey test, $P=.005$).

Table 5. Experience of quick and easy access, sharing of data to reduce doublet registration, and stability of IT^a systems.

Ease of use and staff	Before implementation, mean (SD; range)	After implementation, mean (SD; range)
Quick and easy access to information	2.964 (0.704; 1.0-4.0)	2.959 (0.625; 1.0-4.0)
Medical physician	2.935 (0.6799; 1.0-4.0)	2.816 (0.697; 1.0-4.0)
Medical secretary	3.069 (0.7527; 1.0-4.0)	2.96 (0.611; 1.0-4.0)
Nursing assistant	2.75 (0.5774; 1.0-4.0)	3.0667 (0.594; 1.0-4.0)
Registered nurse	3.011 (0.7219; 1.0-4.0)	3 (0.594; 1.0-4.0)
Sharing of data to reduce doublet registration	2.333 (0.719; 1.0-4.0)	2.3897 (0.705; 1.0-4.0)
Medical physician	2.130 (0.7486; 1.0-4.0)	2.061 (0.8516; 1.0-4.0)
Medical secretary	2.6896 (0.7123; 1.0-4.0)	2.4 (0.5774; 1.0-4.0)
Nursing assistant	2.6875 (0.602; 1.0-4.0)	2.9333 (0.4577; 1.0-4.0)
Registered nurse	2.284 (0.694; 1.0-4.0)	2.456 (0.6227; 1.0-4.0)
Stability of IT systems	2.089 (0.707; 1.0-4.0)	2.359 (0.721; 1.0-4.0)
Medical physician	1.826 (0.7088; 1.0-4.0)	2.245 (0.829; 1.0-4.0)
Medical secretary	2.207 (0.675; 1.0-4.0)	2.32 (0.557; 1.0-4.0)
Nursing assistant	2.25 (0.7746; 1.0-4.0)	2.467 (0.6399; 1.0-4.0)
Registered nurse	2.126 (0.6879; 1.0-4.0)	2.379 (0.7017; 1.0-4.0)

^aIT: information technology.

When looking at the associations between the items for perceived ease of use and staff eHLQ5-7, hypothesis 2.2 was confirmed, as there were highly significant correlations before and after implementation. For the preimplementation measurements, the values were as follows: for the item *ease of access*, staff eHLQ5 ($r=0.2831$; $P<.001$), staff eHLQ6 ($r=0.4385$; $P<.001$), and staff eHLQ7 ($r=0.4164$; $P<.001$); for the item *data is shared between systems to reduce doublet registration*, staff eHLQ5 ($r=0.2055$; $P<.001$), staff eHLQ6 ($r=0.4418$; $P<.001$), and staff eHLQ7 ($r=0.4165$; $P<.001$); and for the item *stability of IT systems*, staff eHLQ5 ($r=0.1753$; $P=.02$), staff eHLQ6 ($r=0.5519$; $P<.001$), and staff eHLQ7 ($r=0.4381$; $P<.001$).

For the postimplementation measurements, the values were as follows: for the item *ease of access*, staff eHLQ5 (*motivated to engage with digital services*; $r=0.3298$; $P<.001$), staff eHLQ6 ($r=0.5237$; $P<.001$), and staff eHLQ7 ($r=0.4759$; $P<.001$); for the item *data is shared between systems to reduce doublet registration*, staff eHLQ5 ($r=0.2763$; $P<.001$), staff eHLQ6 ($r=0.5122$; $P<.001$), and staff eHLQ7 ($r=0.5267$; $P<.001$); and for the item *stability of IT systems*, staff eHLQ5 ($r=0.3402$; $P<.001$), staff eHLQ6 ($r=0.4939$; $P<.001$), and staff eHLQ7 ($r=0.3869$; $P<.001$).

Discussion

Principal Findings

This is the first in-depth examination of medical staff's eHL and perception of ease of use in the transition from a combination of eHealth systems into an integrated EHR. We found that despite a systematic training program, extensive support, and implementation of a coherent EHR, the medical staff's eHL did not change, except for a small decline in

motivation. This is of interest, as the stability of the system is perceived to increase, and the perceived ease of access and the system's ability to share data remain unchanged after the implementation of the EHR.

eHL Scales

Our first hypothesis was an expected increase in all 7 scales of the staff eHLQ based on an increase in knowledge and skills obtained in the implementation process and an increase in the positive experience of using the new system. However, we were unable to confirm this hypothesis.

With respect to eHLQ1 to eHLQ3, relating to personal knowledge and skills, all groups of medical staff had relatively high scores compared with 2 recent studies on medical outpatients and nursing students [19,20]. Regardless of this, only a limited number considered themselves to be experts. Interestingly, there was a positive association between the scores of the eHLQ1 to eHLQ3 scales and the scale regarding how the respondents thought their colleagues would score their user level. This information adds to the evidence for the content validity of the eHLQ1 to eHLQ3 scales. All 4 scales were negatively correlated with the age of respondents after implementation. Interestingly, the association with age was less pronounced before implementation, where only eHLQ3 (*ability to actively engage with digital services*) was associated with age, and Pearson r was lower than that after implementation for the correlation of age and how they thought their colleagues would score their user level. This may indicate that the older part of the respondents experienced less confidence in their self-reported skills as an effect of their experience during the implementation of the new EHR system.

The medical secretaries and nursing assistants scored lower than the medical physicians, which may be related to their prior

training or educational background. Such an association between the level of training or educational background has not yet been observed in relation to eHL; however, further exploration is needed to better understand the possible needs of stratifying digital capacity building.

Before the investigation, we expected that the medical staff's knowledge and skills would increase during the implementation period because of the training and expected higher use of the systems. We were not able to identify such changes as evaluated by the eHLQ scores on scales eHLQ1 to eHLQ3 or in self-reported IT skills described by a colleague. Interestingly, the only change in the staff eHLQ scales was a small decrease in staff eHLQ5 (*motivated to engage with digital services*), indicating that the new EHR system appeared to be less beneficial for users.

As our findings suggest that medical staff report sufficient levels of knowledge and skills but are challenged in relation to how health technology and services are perceived and experienced, we suggest that training should focus on their existing assumptions and prior experiences with the existing EHR.

The lower scores of staff eHLQ5 to eHLQ7 further suggest that the training should focus on how the implementation of the EHR will increase the security and safety of patients, ensure data integration, and support workflows, with data being available to those who need them, including the patients at any time.

Despite the medical physicians having the highest scores in two of the three scales that relate to personal knowledge and skills (eHLQ2 [*understanding of health concepts and language*] and eHLQ3 [*ability to actively engage with digital services*]), they had the lowest scores among the groups of medical staff in three of the four scales relating to their trust in how data are handled (staff eHLQ4, *feeling safe and in control*) and experience with the services (staff eHLQ6 [*access to digital services that work*] and staff eHLQ7 [*digital services that suit individual needs*]).

On the basis of the mean value of eHLQ1 to eHLQ3, our results would suggest that the medical staff's knowledge and skills are not the main issues to be addressed when planning the introduction of a new system. However, when looking at the distribution of scores, it is evident that for eHLQ1 (*using technology to process health information*), 22% (10/46) of medical physicians and 31% (5/16) of nursing assistants scored <2.67. A similar pattern occurred in eHLQ2 (*understanding of health concepts and language*) and eHLQ3 (*ability to actively engage with digital services*), albeit at a lower percentage below the value of 2.67. These results underline the importance of identifying subgroups with low scores across groups of medical staff to address their particular needs in relation to knowledge and skills.

Perceived Ease of Use

Our second hypothesis was that the perceived ease of use, measured by experienced ease of access, ease of data shared between systems to avoid doublet registration, and stability of the system, would increase after the implementation of the EHR system. We only found an increase in the experience of stability of the IT system with the implementation of the new EHR,

which should contribute to a higher perception of ease of use. For nursing assistants and registered nurses, we also found an increase, although not significant, in their score of experience of data being shared between systems to avoid doublet registration, which may be explained by a certain degree of support of their workflows in relation to data. In contrast, medical physicians tended to disagree more than other staff groups with the statement that data were shared between the systems to avoid doublet registrations. This indicates that the system before implementation did not sufficiently support the workflows of medical physicians, and as the medical physicians' degree of disagreement increased after implementation, the new EHR did not have any beneficial effects on their workflows.

Our findings only partly support our hypothesis that the experience relating to the performance of the IT environment would improve within the initial short period of implementation of the first installation of the EHR. The experiences of quick and easy access to relevant information or sharing the data to reduce double registration did not improve overall. As the EHR is provided by one vendor and is anticipated to provide a better experience of coherence and easy access, it is of interest that the medical staff did not experience such an improvement. The new system has many new functions that support quick and easy access. We cannot exclude the possibility that more training and support could have increased the medical staff's capacity to use the system, thereby improving their experience of quick and easy access to information by using macros and tailored interfaces, which the EHR supports.

We also hypothesized that we would be able to identify an association between the staff eHLQ5 to eHLQ7 scales and 1 or more of the 3 items reporting on perceived ease of use: quick and easy access, data being shared between systems to avoid doublet registration, and stability of the IT systems. We found such an association between all 3 staff eHLQ scales and all 3 perceived ease of use items both before and after the implementation.

Confirmation of the hypothesized associations between the staff eHLQ5 to eHLQ7 and the 3 items reporting on perceived ease of use contributes to a better understanding of how eHL, as understood by the eHLF model and measured by staff eHLQ, may intertwine with dimensions from the technology acceptance theory (ie, perceived ease of use and usefulness). User experience is a product of the individual's competence, usability of the user interface, and level of complexity and difficulty of tasks to be solved.

Staff eHLQ5 (*motivated to engage with digital services*) discloses perspectives on the use of health technology, which may relate to a sense of ease of use and usefulness. Staff eHLQ6 (*access to digital services that work*) reports on the experience of data being available whenever needed, and independent of where you are, the data are provided by digital systems that work together. Staff eHLQ7 (*digital services that suit individual needs*) reports on the users' feeling that the digital services suit their needs. In combination, the staff eHLQ5 to eHLQ7 report on this product at a generic level; however, in our study, they were largely influenced by the context of the old EHR or the new EHR system, respectively. User experience, and thereby

the likelihood of adoption [13], is also influenced by the overall perception of how easy and how useful a given technology or system is.

In addition to the respondents' level of eHL, by using the 3 items directly reporting on various aspects of ease of use, we also obtained a more detailed insight into the respondents' specific experience of both the old EHR and the new EHR system and how this relates to their general motivation and experience with health technology.

Implications

Our findings emphasize the need for caution when planning implementations of EHR, as recommended by the literature in this area, such as the studies by McAlearney et al [27] and Boonstra et al [9].

All 4 groups of medical staff had relatively low scores on the staff eHLQ scales, which relates to digital services, and this was most pronounced among medical physicians. If these data had been available to the vendor and the health care organization responsible for the training of the staff, it might have helped them to better address the specific needs of the users; in this case, the medical physicians were characterized by having a high level of self-reported knowledge and skills in relation to data and digital services.

Our findings may also contribute to the understanding of why medical physicians are often resistant to the implementation of EHRs. As pointed out by Boonstra et al [9], this also indicates that an increase in staff capacity with respect to increased digital knowledge and skills may not automatically contribute to an increase in user experience. This is supported by the Monkman and Kushniruk [13] model of adoption, where it is proposed that adoption and a good user experience are both related to the users' eHL and the usability of the systems, as well as the main principle of TAM, which is that the perceived ease of use is a significant factor in facilitating acceptance and adoption [28]. This signifies the importance of tailoring the new EHR's interface and the introduction of system functions according to the specific needs and competences of the medical staff.

The finding that respondents' perception of the ease of access, data sharing to reduce doublet registration, and the stability of the IT systems are influenced by the respondents' overall level of eHL to a large degree suggests that the users' perception of systems is closely linked to not only their competence but also to their general experiences with and confidence in using technology. This knowledge leads us to recommend the identification of staff members with low staff eHLQ scores to better address this particular group specifically and help them during the training to develop or increase self-confidence and self-efficacy in their work with digital health technology.

Limitations

The version of the staff eHLQ used in this study is not the final version and may need further validation. We had to exclude one item from this version as it was not modified to suit the domain to which it belonged. The eHLQ [16] has been thoroughly validated in several languages and appears to be a robust, valid psychometric instrument. The modifications made in the staff

eHLQ do not change the intentions or the significant words of the individual items, and the Cronbach α for the scales demonstrates internal consistency similar to data obtained with the eHLQ. Therefore, we are confident that our results are reliable despite the use of this early version of the staff eHLQ. We also think that the content validity of the staff eHLQ scales is supported by the fact that the staff eHLQ5 and staff eHLQ6 mirror the experience of usefulness of the systems, and eHLQ1 to eHLQ3 is associated with self-reported IT skills, whereas staff eHLQ5 to eHLQ7 is associated with the experience of data being shared in a way that reduces the double entry of data. Another limitation is the lack of administration of the survey after the training but before the implementation of the new EHR system, as we were not able to distinguish between the effect of training and the influence of experiences with the new system, which may affect, for example, the motivation. The reason for this design was a naive approach, where we expected the implementation of the EHR to be beneficial; therefore, we only wanted to focus on the synergy of the new system together with training.

Unfortunately, we were not able to follow up with this after 12 months in the involved department because of restructuring. Therefore, we may have missed effects that would only occur after a longer period of observation, such as 6 or 12 months [29]. We still hope to be able to perform a follow-up later. This is now of particular interest as the vendor in February 2019 has installed a major revision that also has increased interoperability with other national services.

Perspective

The digital competence of the medical staff may vary among countries and regions and may therefore be addressed differently when a vendor or organization introduces a new EHR system. The staff eHLQ may be used to better understand the particular needs of medical staff groups, which should be addressed.

In addition, staff eHLQ6 may have an important role in settings where the EHR is not only used for documentation of hospital activities but also for primary care activities, and data are expected to be available for all actors at any time. However, this requires further investigation.

The association between the level of eHL and indicators of how the respondents perceive the performance of the system calls for further research on whether ≥ 1 of the staff eHLQ scales (ie, eHLQ5-7) can be used as predictors for users' acceptance of technology in health care settings.

Conclusions

The staff eHLQ may be a good candidate for monitoring the medical staff's response to their training during the implementation of a new EHR system. It may also inform those responsible for the implementation whether the process is not going according to plan, with respect to the staff's knowledge, skills, trust in security, motivation, and experience of a coherent system that suits the needs and supports the workflow and data availability.

Overall, this new insight in the presented case could have been helpful for the organization that led the implementation of the

EHR and helped them to understand how the training should focus on how to (1) make use of the new functionality, (2) inform about the changes in workflow, and (3) make sense of the transition and thereby focus less on digital competence. It should be noted that the lower scores of staff eHLQ5 to eHLQ7, as found in all groups of medical staff, may also be because of

problems with the functionality of the EHR as it was the first installation.

This calls for both the vendors in their design and the health care organizations in their procurement to pay more attention to these areas in the implementation process.

Acknowledgments

The authors would like to thank the medical and administrative staff at Department C, Herlev-Gentofte Hospital, for their support and participation in the study. The authors would also like to thank Ms Astrid Laura Dam Jensen for her assistance with the editing of this paper and preparation of the manuscript for submission.

This project has not received any external funding.

Conflicts of Interest

None declared.

References

1. The Topol Review Preparing the Healthcare Workforce to Deliver the Digital Future. England, United Kingdom: National Health Service; 2019.
2. Greenhalgh T, Potts H, Wong G, Bark P, Swinglehurst D. Tensions and paradoxes in electronic patient record research: a systematic literature review using the meta-narrative method. *Milbank Q* 2009 Dec;87(4):729-788 [FREE Full text] [doi: [10.1111/j.1468-0009.2009.00578.x](https://doi.org/10.1111/j.1468-0009.2009.00578.x)] [Medline: [20021585](https://pubmed.ncbi.nlm.nih.gov/20021585/)]
3. Peterson C, Hamilton C, Hasvold P. From Innovation to Implementation - EHealth in the WHO European Region. Copenhagen, Denmark: WHO Regional Office for Europe; 2016.
4. Kuek A, Hakkennes S. Healthcare staff digital literacy levels and their attitudes towards information systems. *Health Informatics J* 2020 Mar 15;26(1):592-612 [FREE Full text] [doi: [10.1177/1460458219839613](https://doi.org/10.1177/1460458219839613)] [Medline: [30983476](https://pubmed.ncbi.nlm.nih.gov/30983476/)]
5. WHO guideline. World Health Organization. URL: <https://tinyurl.com/2p8pdxnz> [accessed 2022-03-26]
6. Assessing the Impact of Digital Transformation of Health Services. Brussels: European Commission; 2019.
7. Expert Panel on Effective Ways of Investing in Health (EXPH), Report on Access to Health Services in the European Union. Brussels: European Union; 2016.
8. Tsai CH, Eghdam A, Davoody N, Wright G, Flowerday S, Koch S. Effects of electronic health record implementation and barriers to adoption and use: a scoping review and qualitative analysis of the content. *Life (Basel)* 2020 Dec 04;10(12):327 [FREE Full text] [doi: [10.3390/life10120327](https://doi.org/10.3390/life10120327)] [Medline: [33291615](https://pubmed.ncbi.nlm.nih.gov/33291615/)]
9. Boonstra A, Versluis A, Vos JF. Implementing electronic health records in hospitals: a systematic literature review. *BMC Health Serv Res* 2014 Sep 04;14(1):370 [FREE Full text] [doi: [10.1186/1472-6963-14-370](https://doi.org/10.1186/1472-6963-14-370)] [Medline: [25190184](https://pubmed.ncbi.nlm.nih.gov/25190184/)]
10. Hennemann S, Beutel ME, Zwerenz R. Ready for eHealth? Health professionals' acceptance and adoption of eHealth interventions in inpatient routine care. *J Health Commun* 2017 Mar;22(3):274-284. [doi: [10.1080/10810730.2017.1284286](https://doi.org/10.1080/10810730.2017.1284286)] [Medline: [28248626](https://pubmed.ncbi.nlm.nih.gov/28248626/)]
11. Davis FD. Perceived usefulness, perceived ease of use, and user acceptance of information technology. *MIS Q* 1989 Sep;13(3):319. [doi: [10.2307/249008](https://doi.org/10.2307/249008)]
12. Venkatesh V, Morris MG, Davis GB, Davis FD. User acceptance of information technology: toward a unified view. *MIS Q* 2003;27(3):425-478. [doi: [10.2307/30036540](https://doi.org/10.2307/30036540)]
13. Monkman H, Kushniruk AW. The consumer health information system adoption model. *Stud Health Technol Inform* 2015;218:26-31. [Medline: [26262522](https://pubmed.ncbi.nlm.nih.gov/26262522/)]
14. Norman CD, Skinner HA. eHealth literacy: essential skills for consumer health in a networked world. *J Med Internet Res* 2006 Jun 16;8(2):e9 [FREE Full text] [doi: [10.2196/jmir.8.2.e9](https://doi.org/10.2196/jmir.8.2.e9)] [Medline: [16867972](https://pubmed.ncbi.nlm.nih.gov/16867972/)]
15. Norgaard O, Furstrand D, Klokke L, Karnoe A, Batterham R, Kayser L, et al. The e-health literacy framework: a conceptual framework for characterizing e-health users and their interaction with e-health systems. *Knowl Manag E-Learn Int J* 2015;7(4):540. [doi: [10.34105/j.kmel.2015.07.035](https://doi.org/10.34105/j.kmel.2015.07.035)]
16. Kayser L, Karnoe A, Furstrand D, Batterham R, Christensen KB, Elsworth G, et al. A multidimensional tool based on the ehealth literacy framework: development and initial validity testing of the eHealth literacy questionnaire (eHLQ). *J Med Internet Res* 2018 Feb 12;20(2):e36 [FREE Full text] [doi: [10.2196/jmir.8371](https://doi.org/10.2196/jmir.8371)] [Medline: [29434011](https://pubmed.ncbi.nlm.nih.gov/29434011/)]
17. Sounderajah V, Clarke J, Yalamanchili S, Acharya A, Markar SR, Ashrafian H, et al. A national survey assessing public readiness for digital health strategies against COVID-19 within the United Kingdom. *Sci Rep* 2021 Mar 16;11(1):5958 [FREE Full text] [doi: [10.1038/s41598-021-85514-w](https://doi.org/10.1038/s41598-021-85514-w)] [Medline: [33727655](https://pubmed.ncbi.nlm.nih.gov/33727655/)]

18. Thies K, Anderson D, Cramer B. Lack of adoption of a mobile app to support patient self-management of diabetes and hypertension in a federally qualified health center: interview analysis of staff and patients in a failed randomized trial. *JMIR Hum Factors* 2017 Oct 03;4(4):e24 [FREE Full text] [doi: [10.2196/humanfactors.7709](https://doi.org/10.2196/humanfactors.7709)] [Medline: [28974481](https://pubmed.ncbi.nlm.nih.gov/28974481/)]
19. Holt KA, Karnoe A, Overgaard D, Nielsen SE, Kayser L, Røder ME, et al. Differences in the level of electronic health literacy between users and nonusers of digital health services: an exploratory survey of a group of medical outpatients. *Interact J Med Res* 2019 Apr 05;8(2):e8423 [FREE Full text] [doi: [10.2196/ijmr.8423](https://doi.org/10.2196/ijmr.8423)] [Medline: [30950809](https://pubmed.ncbi.nlm.nih.gov/30950809/)]
20. Holt KA, Overgaard D, Engel LV, Kayser L. Health literacy, digital literacy and eHealth literacy in Danish nursing students at entry and graduate level: a cross sectional study. *BMC Nurs* 2020 Apr 10;19(1):22 [FREE Full text] [doi: [10.1186/s12912-020-00418-w](https://doi.org/10.1186/s12912-020-00418-w)] [Medline: [32308559](https://pubmed.ncbi.nlm.nih.gov/32308559/)]
21. Thorsen IK, Rossen S, Glümer C, Midtgaard J, Ried-Larsen M, Kayser L. Health technology readiness profiles among Danish individuals with type 2 diabetes: cross-sectional study. *J Med Internet Res* 2020 Sep 15;22(9):e21195 [FREE Full text] [doi: [10.2196/21195](https://doi.org/10.2196/21195)] [Medline: [32930669](https://pubmed.ncbi.nlm.nih.gov/32930669/)]
22. Rossen S, Kayser L, Vibe-Petersen J, Christensen JF, Ried-Larsen M. Cancer survivors' receptiveness to digital technology-supported physical rehabilitation and the implications for design: qualitative study. *J Med Internet Res* 2020 Aug 05;22(8):e15335 [FREE Full text] [doi: [10.2196/15335](https://doi.org/10.2196/15335)] [Medline: [32755892](https://pubmed.ncbi.nlm.nih.gov/32755892/)]
23. Terp R, Kayser L, Lindhardt T. "It is not rocket science.." - Older peoples' understanding of nutrition - A qualitative study. *Appetite* 2021 Jan 01;156:104854. [doi: [10.1016/j.appet.2020.104854](https://doi.org/10.1016/j.appet.2020.104854)] [Medline: [32871203](https://pubmed.ncbi.nlm.nih.gov/32871203/)]
24. Region Hovedstaden, Region Sjælland. Intro til Sundhedsplatformen - Få overblik: Hvad betyder Sundhedsplatformen for patienter og personale?. Denmark: Sundheds Platformen; Oct 11, 2017.
25. International Standard Classification of Education ISCED 2011. Montreal, Canada: UNESCO Institute for Statistics; 2012.
26. Undersøgelse af klinisk anvendelse af sundheds-it-systemer 2014. Dansk Center for Sundhedsinformatik. URL: https://www.dachi.aau.dk/digitalAssets/131/131389_15-1-undersogelse-klinisk-it-anvendelse-2014.pdf [accessed 2022-03-26]
27. McAlearney AS, Robbins J, Kowalczyk N, Chisolm DJ, Song PH. The role of cognitive and learning theories in supporting successful EHR system implementation training: a qualitative study. *Med Care Res Rev* 2012 Jun 26;69(3):294-315. [doi: [10.1177/1077558711436348](https://doi.org/10.1177/1077558711436348)] [Medline: [22451617](https://pubmed.ncbi.nlm.nih.gov/22451617/)]
28. Venkatesh V, Bala H. Technology acceptance model 3 and a research agenda on interventions. *Decision Sci* 2008 May;39(2):273-315. [doi: [10.1111/j.1540-5915.2008.00192.x](https://doi.org/10.1111/j.1540-5915.2008.00192.x)]
29. El-Kareh R, Gandhi TK, Poon EG, Newmark LP, Ungar J, Lipsitz S, et al. Trends in primary care clinician perceptions of a new electronic health record. *J Gen Intern Med* 2009 Apr;24(4):464-468 [FREE Full text] [doi: [10.1007/s11606-009-0906-z](https://doi.org/10.1007/s11606-009-0906-z)] [Medline: [19156468](https://pubmed.ncbi.nlm.nih.gov/19156468/)]

Abbreviations

- eHL:** eHealth literacy
- eHLF:** eHealth Literacy Framework
- eHLQ:** eHealth Literacy Questionnaire
- EHR:** electronic health record
- ISCED:** International Standard Classification of Education
- IT:** information technology
- RQ:** research question
- TAM:** Technology Acceptance Model
- UTAUT:** Unified Theory of Acceptance and Use of Technology

Edited by A Kushniruk; submitted 20.04.21; peer-reviewed by N Werts, E Brainin, E Neter, JR Bautista; comments to author 08.07.21; revised version received 22.11.21; accepted 20.01.22; published 29.04.22.

Please cite as:

Kayser L, Karnoe A, Duminski E, Jakobsen S, Terp R, Dansholm S, Roeder M, From G
Health Professionals' eHealth Literacy and System Experience Before and 3 Months After the Implementation of an Electronic Health Record System: Longitudinal Study
JMIR Hum Factors 2022;9(2):e29780
URL: <https://humanfactors.jmir.org/2022/2/e29780>
doi: [10.2196/29780](https://doi.org/10.2196/29780)
PMID: [35486414](https://pubmed.ncbi.nlm.nih.gov/35486414/)

©Lars Kayser, Astrid Karnoe, Emily Duminski, Svend Jakobsen, Rikke Terp, Susanne Dansholm, Michael Roeder, Gustav From. Originally published in *JMIR Human Factors* (<https://humanfactors.jmir.org/>), 29.04.2022. This is an open-access article distributed

under the terms of the Creative Commons Attribution License (<https://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Human Factors, is properly cited. The complete bibliographic information, a link to the original publication on <https://humanfactors.jmir.org>, as well as this copyright and license information must be included.

Original Paper

A Web-Based Health Application to Translate Nutrition Therapy for Cardiovascular Risk Reduction in Primary Care (PortfolioDiet.app): Quality Improvement and Usability Testing Study

Meaghan E Kavanagh^{1,2,3}, MSc; Laura Chiavaroli^{1,2,3}, MSc, PhD; Andrea J Glenn^{1,2,3,4}, RD, PhD; Genevieve Heijmans³; Shannan M Grant^{5,6,7}, Pdt, RD, MSc, PhD; Chi-Ming Chow⁸, MSc, MD; Robert G Josse^{1,2,3,9,10}, MD; Vasanti S Malik^{1,4}, PhD; William Watson¹¹, MD; Aisha Lofters^{12,13,14}, MD; Candice Holmes¹¹, MD; Julia Rackal¹¹, MD; Kristie Srichaikul^{2,3}, MSc, MD; Diana Sherifali¹⁵, RN, BScN, PhD; Erna Snelgrove-Clarke¹⁶, RN, PhD; Jacob A Udell^{17,18,19,20,21,22}, MD; Peter Juni^{9,18,19}, MD; Gillian L Booth^{19,20,21,23}, MSc, MD; Michael E Farkouh²², MD, MSc; Lawrence A Leiter^{1,2,3,9,10,21}, MD; Cyril W C Kendall^{1,2,3,24}, PhD; David J A Jenkins^{1,2,3,9,10,21}, MD, PhD; John L Sievenpiper^{1,2,3,9,10,21}, MD, PhD

¹Department of Nutritional Sciences, Temerty Faculty of Medicine, University of Toronto, Toronto, ON, Canada

²Clinical Nutrition and Risk Factor Modification Center, St. Michael's Hospital, Unity Health Toronto, Toronto, ON, Canada

³Toronto 3D Knowledge Synthesis and Clinical Trials Unit, St. Michael's Hospital, Unity Health Toronto, Toronto, ON, Canada

⁴Department of Nutrition, Harvard T.H. Chan School of Public Health, Boston, MA, United States

⁵Department of Applied Human Nutrition, Mount Saint Vincent University, Halifax, NS, Canada

⁶Departments of Pediatrics and Obstetrics and Gynaecology, IWK Health, Halifax, NS, Canada

⁷Department of Obstetrics and Gynaecology, Faculty of Medicine, Dalhousie University, Halifax, NS, Canada

⁸Division of Cardiology, St. Michael's Hospital, Unity Health Toronto, Toronto, ON, Canada

⁹Li Ka Shing Knowledge Institute, St. Michael's Hospital, Unity Health Toronto, Toronto, ON, Canada

¹⁰Division of Endocrinology and Metabolism, St. Michael's Hospital, Unity Health Toronto, Toronto, ON, Canada

¹¹Department of Family and Community Medicine, St. Michael's Hospital, Unity Health Toronto, Toronto, ON, Canada

¹²Department of Family and Community Medicine, Temerty Faculty of Medicine, University of Toronto, Toronto, ON, Canada

¹³Family Practice Health Centre, Women's College Hospital, Toronto, ON, Canada

¹⁴Peter Gilgan Centre for Women's Cancers, Women's College Hospital, Toronto, ON, Canada

¹⁵School of Nursing, Faculty of Health Sciences, McMaster University, Hamilton, ON, Canada

¹⁶School of Nursing, Faculty of Health Sciences, Queen's University, Kingston, ON, Canada

¹⁷Women's College Research Institute and Cardiovascular Division, Department of Medicine, Women's College Hospital, University of Toronto, Toronto, ON, Canada

¹⁸Applied Health Research Centre, Li Ka Shing Knowledge Institute of St. Michael's Hospital, Department of Medicine, University of Toronto, Toronto, ON, Canada

¹⁹Institute for Health Policy, Management, and Evaluation, Dalla Lana School of Public Health, University of Toronto, Toronto, ON, Canada

²⁰ICES, Toronto, ON, Canada

²¹Department of Medicine, Temerty Faculty of Medicine, University of Toronto, Toronto, ON, Canada

²²Peter Munk Cardiac Centre and the Heart and Stroke Richard Lewar Centre, University of Toronto, Toronto, ON, Canada

²³MAP Centre for Urban Health Solutions, Li Ka Shing Knowledge Institute, St. Michael's Hospital, Toronto, ON, Canada

²⁴College of Pharmacy and Nutrition, University of Saskatchewan, Saskatoon, SK, Canada

Corresponding Author:

John L Sievenpiper, MD, PhD

Department of Nutritional Sciences

Temerty Faculty of Medicine

University of Toronto

5th floor, room 5334, Medical Sciences Building

1 King's College Circle

Toronto, ON, M5S 1A8

Canada

Phone: 1 416 867 3732

Fax: 1 416 867 7495

Email: john.sievenpiper@utoronto.ca

Abstract

Background: The Portfolio Diet, or Dietary Portfolio, is a therapeutic dietary pattern that combines cholesterol-lowering foods to manage dyslipidemia for the prevention of cardiovascular disease. To translate the Portfolio Diet for primary care, we developed the PortfolioDiet.app as a patient and physician educational and engagement tool for PCs and smartphones. The PortfolioDiet.app is currently being used as an add-on therapy to the standard of care (usual care) for the prevention of cardiovascular disease in primary care. To enhance the adoption of this tool, it is important to ensure that the PortfolioDiet.app meets the needs of its target end users.

Objective: The main objective of this project is to undertake user testing to inform modifications to the PortfolioDiet.app as part of ongoing engagement in quality improvement (QI).

Methods: We undertook a 2-phase QI project from February 2021 to September 2021. We recruited users by convenience sampling. Users included patients, family physicians, and dietitians, as well as nutrition and medical students. For both phases, users were asked to use the PortfolioDiet.app daily for 7 days. In phase 1, a mixed-form questionnaire was administered to evaluate the users' perceived acceptability, knowledge acquisition, and engagement with the PortfolioDiet.app. The questionnaire collected both quantitative and qualitative data, including 2 open-ended questions. The responses were used to inform modifications to the PortfolioDiet.app. In phase 2, the System Usability Scale was used to assess the usability of the updated PortfolioDiet.app, with a score higher than 70 being considered acceptable.

Results: A total of 30 and 19 users were recruited for phase 1 and phase 2, respectively. In phase 1, the PortfolioDiet.app increased users' perceived knowledge of the Portfolio Diet and influenced their perceived food choices. Limitations identified by users included challenges navigating to resources and profile settings, limited information on plant sterols, inaccuracies in points, timed-logout frustration, request for step-by-step pop-up windows, and request for a mobile app version; when looking at positive feedback, the recipe section was the most commonly praised feature. Between the project phases, 6 modifications were made to the PortfolioDiet.app to incorporate and address user feedback. At phase 2, the average System Usability Scale score was 85.39 (SD 11.47), with 100 being the best possible.

Conclusions: By undertaking user testing of the PortfolioDiet.app, its limitations and strengths were able to be identified, informing modifications to the application, which resulted in a clinical tool that better meets users' needs. The PortfolioDiet.app educates users on the Portfolio Diet and is considered acceptable by users. Although further refinements to the PortfolioDiet.app will continue to be made before its evaluation in a clinical trial, the result of this QI project is an improved clinical tool.

(*JMIR Hum Factors* 2022;9(2):e34704) doi:[10.2196/34704](https://doi.org/10.2196/34704)

KEYWORDS

portfolio diet; dietary portfolio; nutrition therapy; dietary application; eHealth; usability testing; quality improvement; mobile phone

Introduction

Background

The Portfolio Diet, or Dietary Portfolio, is a plant-based dietary pattern of cholesterol-lowering foods that has demonstrated *drug-like* reductions in low-density lipoprotein cholesterol (LDL-C) and other cardiovascular risk factors [1,2]. In a metabolically controlled study, the Portfolio Diet was shown to result in the same LDL-C reduction (approximately 30%) as lovastatin therapy, the first statin to be widely used [1]. In a recent systematic review and meta-analysis, these *drug-like* reductions in LDL-C were confirmed and further benefits were also found on other aspects of the lipid profile (non-high-density lipoprotein cholesterol, apolipoprotein B, and triglycerides), blood pressure, inflammatory markers, and estimated 10-year Framingham risk score compared with a National Cholesterol Education Program Step 2 diet alone [3]. This evidence has led to the recognition of the Portfolio Diet as a therapy for

cardiovascular disease management from major international clinical practice guidelines, including the Canadian Cardiovascular Society [4,5], Diabetes Canada [6], Obesity Canada [7], Canadian Cardiovascular Harmonized National Guidelines Endeavour [8], Heart UK [9], European Atherosclerosis Society [10], and the American College of Cardiology and American Heart Association guidelines [11]. Although the Portfolio Diet is recognized by clinical practice guidelines as a preventive nutrition therapy for cardiovascular disease, implementation in clinical practice is limited. Traditionally, nutrition therapy involves multiple face-to-face sessions over an extended length of time with trained personnel. However, many health care providers cite a lack of education, educational materials, and time to counsel their patients on nutrition [12,13]. Advancements in technology may be able to circumvent these issues and expand access to nutrition therapies for patients.

Several studies have shown that health apps can promote positive behavior change and improve related health outcomes. Block et al [14] found that a fully automated intervention targeting nutritional and physical activity behaviors in individuals with prediabetes improved glycemic control and Framingham diabetes risk score over 6 months compared with the waitlist control. In a meta-analysis of 47 randomized controlled trials, Beishuizen et al [15] found that web-based interventions in primary care settings improved risk factors for cardiovascular disease compared with standard of care alone. Thus, health apps can provide an alternative and complementary approach to delivery of preventive nutrition therapy within the limits of primary care, where the shift to remote care during the COVID-19 pandemic has further highlighted the need for evidence-based health apps [16].

Therefore, to translate the current clinical practice guidelines for nutrition therapy for dyslipidemia, we developed a web-based application, the PortfolioDiet.app. The application was developed by an interdisciplinary team of clinical nutrition experts, registered dietitians, cardiologists, and software architects, as well as patient, physician, and dietitian advisory committees. The collaboration with knowledge users throughout the development and testing process is the central premise of the integrated knowledge translation (iKT) approach [17]. The PortfolioDiet.app is currently being used as an optional add-on therapy to the standard of care (usual care) for primary and secondary prevention of cardiovascular disease at St Michael's Hospital, Toronto, Ontario, Canada. As part of this iKT approach to enhance the adoption of the Portfolio Diet, it is important to ensure that the PortfolioDiet.app meets the needs of its end users. The population of end users for the application includes adult patients at risk for cardiovascular disease and clinical staff who may wish to learn more about the diet or want to recommend the PortfolioDiet.app to their patients. By undertaking user-centered evaluations, the needs of the target population can be identified, leading to improved uptake of the application.

Quality Improvement Initiatives

Quality improvement (QI) initiatives offer an opportunity to optimize and test current clinical tools and are a proven method to improve patient care [18]. These initiatives are especially important when the clinical tool is an app because usability problems have been identified as a major obstacle in the adoption of health apps and have been associated with attrition [19,20]. By performing usability testing of health apps, problems related to ease of use can be identified before undertaking costly trials. Although regarded by many as an essential step in app development, usability testing of nutrition apps is less common in the literature, possibly leading to low user engagement and loss of effectiveness over time [21]. In a recent systematic review by König et al [22], usability was the most frequently identified barrier by participants for nutrition apps, underpinning the importance of usability testing in the development of nutrition therapy apps. This paper provides a description of our user testing approach to help inform research groups seeking to improve similar apps. Therefore, the objective of this project is to undertake and describe user testing to inform modifications to the PortfolioDiet.app as part of ongoing engagement in QI.

Methods

System Intervention

The PortfolioDiet.app is based on a nutrition therapy to manage dyslipidemia, the Portfolio Diet, that was demonstrated to be effective in individuals with hyperlipidemia [1,2]. The PortfolioDiet.app is a freely available web-based application that can be accessed on any smartphone or PC [23]. A web-based platform was chosen as the initial form to ensure that the application was accessible to patients. Although most Canadians have home internet or smartphone access (94% and 86%, respectively, in 2017) [24], having a web-based platform allows those patients who do not have home internet or a smartphone to access the application through public computers such those as in libraries. Ensuring accessibility was especially important, given the inner city community that St Michael's Hospital serves. The PortfolioDiet.app is automated and patient facing. The application contains a variety of personalized elements to enhance and sustain patient education and engagement based on a 25-point Portfolio Diet score. These include elements preferred by health app users: an interaction-enabled dashboard, learning resources, gamification components, nudging, and so on [25]. The dashboard presents various summary statistics on adherence, such as total score, individual diet component score, and a 30-day score trend (Multimedia Appendix 1 shows screenshots depicting the various features on the dashboard of the application). The learning resources in the application include the Portfolio Diet infographic, recipes, tip sheets, and educational videos. The infographic provides a visual of the Portfolio Diet and its health benefits (Multimedia Appendix 2). The recipes were developed by registered dietitians according to the Portfolio Diet's targets. The gamification components include star rewards, weekly quiz questions about the Portfolio Diet, and the Portfolio Diet score leaderboard. Users gain star rewards for each log-in of the day and for completion of weekly quiz questions.

Design

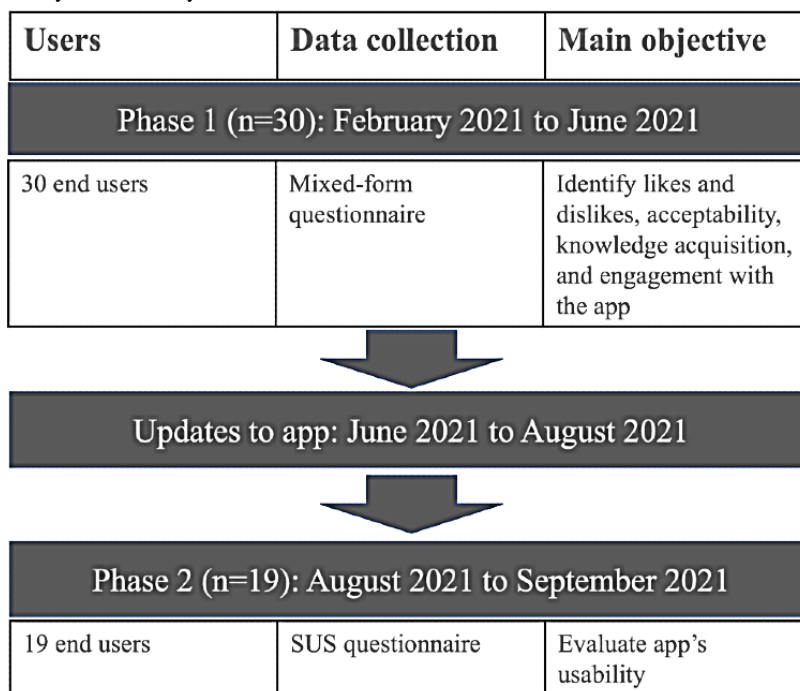
We performed a 2-phase QI project from February 2021 to September 2021 (Figure 1). Adult users from a number of areas were invited by email to participate in the testing of the application. Selective convenience sampling was used to generate a varied sample of previously identified end-user groups: patients with hyperlipidemia, family physicians and registered dietitians, the general public, and nutrition and medical students. For both phases, users were provided with the PortfolioDiet.app link and an instructional guide (Multimedia Appendix 3 shows example pages from the instructional guide) and asked to use the PortfolioDiet.app daily for 7 days. As the application is intended to be used over a long-term duration, a 7-day time frame was chosen to ensure that the users had sufficient time to experience each PortfolioDiet.app feature, such as the personalized weekly email reminders and accumulation of their daily scores displayed on the 30-day graph.

For phase 1, users were sent an email at the end of the 7 days, asking them to complete a mixed-form questionnaire and return it by email (Multimedia Appendix 4). The questionnaire was

developed with experts in knowledge uptake evaluation. The purpose of the questionnaire was to evaluate the user’s perceived acceptability, knowledge acquisition, and engagement with the PortfolioDiet.app. The questionnaire collected both quantitative and qualitative data, including 2 open-ended questions. A mixed-form questionnaire allows for a more comprehensive collection of data on views and feedback from end users [26]. Open-ended questions were included to provide users an opportunity to identify strengths and limitations. In phase 2, the usability of the PortfolioDiet.app was measured using the System Usability Scale (SUS; Multimedia Appendix 5). The SUS is a validated usability questionnaire that has been used in clinical settings to assess the usability of various systems and tools [27,28]. The SUS includes 10 statements rated on a 5-point Likert scale. The Likert scale is a psychometric scale often used in psychology questionnaires and frequently applied in health, nutrition, and foods research as well as QI to assess the acceptability of systems and tools. These scales are often used to assess personality, attitudes, and behaviors.

The application users were asked to indicate their age range (<40 years, 40-60 years, >60 years) because age has been previously identified as an important covariate when assessing usability and is inversely correlated with the SUS score, whereas other characteristics such as gender have not [29]. Other than age range, no other demographic information from users was collected. During the QI project, the core team held weekly meetings to coordinate the application development process. User feedback from phase 1 was discussed within the research team during these weekly meetings. Modifications and updates to address user feedback were implemented in the PortfolioDiet.app and its supporting material before the initiation of phase 2. Although sample sizes of n=5 have previously been deemed acceptable for usability testing [30], a sample of at least 25 was decided upon to ensure a high level of problem detection [31]; therefore, with an anticipated response rate of 80% [32], a total of 30 users were invited. All 30 invitees accepted.

Figure 1. Project overview. SUS: System Usability Scale.



Data Analysis

All open-ended responses from the questionnaire were collected and sorted manually into either limitations or strengths. Common comments (reported by ≥2 users) were identified and summarized. Representative quotations of common comments were included to improve the credibility of the findings, an approach recommended by Graneheim and Lundman [33]. All quantitative data were grouped and summarized as totals. A summary usability score was calculated (range 0-100) using the standard score conversion procedure for the SUS [27], with a score higher than 70 being considered acceptable [29]. Data were expressed as mean (SD).

Ethics Approval

This project was formally reviewed by institutional authorities at Unity Health Toronto and deemed to require neither research ethics board approval nor written informed consent from participants.

Results

Phase 1

For phase 1, a total of 30 users provided feedback from February 2021 to June 2021, with 20 (66%) users aged <40 years, 6 (20%) aged 40-60 years, and 4 (13%) aged >60 years. The response rate for the questionnaire was 100% (30/30). Table S1 in Multimedia Appendix 6 presents the results of the quantitative responses of phase 1. Of the 30 users, 29 (97%) said the PortfolioDiet.app increased their knowledge of the Portfolio

Diet, demonstrating that perceived knowledge acquisition was high. Most of the users reported that the application influenced or changed their food choices (24/30, 80%) and that they would use the application daily (20/30, 67%) or weekly (9/30, 30%), demonstrating a high level of engagement with the application. In addition, users ranked the infographic and the tip sheets as the first and second highest features that helped them learn about the Portfolio Diet. Users ranked the star rewards (a gamification component) and recipes as the first and second highest features that supported their interest and engagement in using the application. Most of the users responded that the application was easy to use (26/30, 87%) and it was easy to navigate

between the applications functions (28/30, 93%), demonstrating acceptability. Common comments from ≥ 2 users are summarized as representative quotations in [Textbox 1](#) (a full report of all comments can be found in Table S2 in [Multimedia Appendix 6](#)).

Feedback from phase 1 revealed several opportunities for improvement of the PortfolioDiet.app and its content. The user feedback was reviewed by the team during weekly meetings, and modifications to the second prototype of the application were made to address the common comments. Subsequently, usability of the updated application prototype was assessed in phase 2.

Textbox 1. Qualitative data from users in phase 1.

Representative quotations of feedback on the PortfolioDiet.app after using it for 7 days.

1. Limitations and suggestions for improvement

- “I got more familiar with the food items in each category. I might have learned more but I didn’t realize at first that there was anything important in the ‘Learn’ section.”
- “I found some portions to be very large.”
- “...one thing I did not enjoy was the lack of information about plant sterols and where to find/purchase these.”
- “The app did not accurately record my average scores.”
- “It would be nice not to have to log in each and every time, if the app could remember my login info.”
- “It would be helpful to possibly add a video or a step-by-step guide that pops up when you first enter the app. Otherwise, it felt like I had to search for the diet outline and recipes myself.”
- “Would prefer an actual app, and not doing it via web browser.”

2. Positives and strengths

- “I learned a lot about which foods are part of the diet as well as quantities needed for one serving”
- “The resources gave good summaries of the Portfolio Diet. I was unaware of the Diet prior to beginning using the app so it was a good introduction. The tip sheets and recipes were very helpful.”
- “I found the front page most useful by allowing me to see where I am not meeting the daily targets, and where and what I still need to eat for the day.”
- “I loved how easy it was to enter information into the app and the progress bar really helped me visualize my progress.”
- “The recipe booklet offered many great and creative meal ideas, and I can personally say I have used it since, and will continue to use it moving forward.”

App Updates Based on Themes

Navigation to Resources: Theme 1

Users reported navigation challenges with the phase 1 application prototype. For example, some users were unsure how to initially navigate through the PortfolioDiet.app to find resources. Although a PDF instructional guide was provided to all users, this may have not been the most suitable format for communicating with all user types. Therefore, several short videos were created to help supplement the PDF instructional guide, resulting in 9 tutorial videos lasting from 1 minute to 3 minutes to familiarize users with the PortfolioDiet.app and its functions.

Navigation to Settings: Theme 2

Multiple users noted concerns with the food portions within the application; a user commented: “I found some portions to be very large.” The application automatically starts all users on

the 2000 kcal per day diet. Although instructions for users on changing calorie targets were provided through the PDF instructional guide, it may have not been intuitive to users that the fruit icon at the top of the home page would lead them to their account settings. Therefore, to help users navigate to their account settings, the fruit icon was changed to an *Account* button ([Multimedia Appendix 7](#) shows the screenshots of the application changes). In addition, a short instructional video was created that explained to users how to correctly select their appropriate kcal target per day in the account settings of the application.

Plant Sterol Familiarity: Theme 3

Users expressed limited knowledge of plant sterols. As plant sterols are 1 of 5 major components of the Portfolio Diet, it is important that patients feel informed and comfortable incorporating plant sterols into their diet. It was decided that the creation of an evidence-based educational resource was

critical to helping communicate the health benefits of plant sterols to patients and clinical staff. Therefore, a plant sterol tip sheet was developed and added to the updated version of the application ([Multimedia Appendix 7](#)).

Saving Issue Leading to Point Inaccuracy: Theme 4

Of the 30 users, 2 (7%) expressed concerns regarding inaccuracies in their 25-point Portfolio Diet score calculated by the application. An investigation with the application development team determined that after 45 minutes the application was not connecting with the server and not saving food entries for some users. To address this, a logout notification was added to inform users when to refresh and log back into the application ([Multimedia Appendix 7](#)).

Logout Frustration: Theme 5

Users expressed frustration with the application automatically logging them out after 45 minutes of inactivity. To help reduce user frustration, the automatic logout was extended to 21 hours as a balance between user experience and personal health data security.

Opportunities for Future Improvements: Themes 6 and 7

Although many improvements to the PortfolioDiet.app were made based on phase 1 feedback, certain user feedback remained challenging to address in the short term ([Textbox 1](#): themes 6 and 7). Users suggested the addition of pop-up windows to help with initial navigation to important areas. Although the QI team agreed with the benefits of pop-ups, this proved challenging to implement and was considered lower priority than other key application issues identified by users. In addition, users expressed interest in, or preference for, a mobile app over the current web-based platform. A web-based platform was chosen as the initial form to ensure accessibility of the application. Future work to enhance the adoption of this tool will include the development of an iOS app and an Android app for mobile use as well as the integration of pop-ups to further engage participants with features and resources.

Recipes: Theme 12

The enjoyment of the recipes was the most commonly praised feature by users and was the second favorite application feature supporting engagement with the PortfolioDiet.app (Table S1 in [Multimedia Appendix 6](#)). Therefore, the recipe bank was expanded from 53 to 70 recipes and culinary students were engaged in this work to expand the cultural diversity of the recipes. To enhance usability, the downloadable PDF recipe book was converted into a filterable recipe webpage, allowing users to filter recipes by each of the Portfolio Diet categories. Recipes were also made filterable by type of meal (eg, breakfast, lunch, dinner, and snack), preparation difficulty level (eg, beginner and intermediate), and preparation time (eg, quick; [Multimedia Appendix 7](#)).

Phase 2

In phase 2, a total of 19 users completed the SUS from August 2021 to September 2021, with 11 (58%) users aged <40 years, 5 (26%) aged 40-60 years, and 3 (16%) aged >60 years. The response rate was 79% (19/24). Nearly half of the participants

(9/19, 47%) were new to using the application. The rest were previously users in the phase 1 testing who were reapproached and asked to again use the updated application for 7 days. The participating users gave the application a mean SUS score of 85.39 (SD 11.47). Full responses to the individual SUS items are shown in Table S3 in [Multimedia Appendix 6](#). Examination of the responses to the individual SUS items showed that most users thought that they would not need the support of a technical person to use the application (average rating of 1.11, SD 0.32, out of 5, where 1=strongly disagree and 5=strongly agree), they thought that the application was easy to use (4.47, SD 0.84), they believed that most people would learn to use the application very quickly (4.68, SD 0.67), and they felt confident using the application (4.32, SD 0.76). There were 2 questions where, although most of the users agreed that they would use the application frequently (3.58, SD 0.90) and that the various functions in the application were well integrated (3.79, SD 1.32), these scores averaged closer to a neutral rating; therefore, updates to further improve application engagement and application function integration will be a focus during the next application revision.

Discussion

Principal Findings

The result of this QI project is a clinical tool that better meets the needs of end users. Through this 2-phase QI project, user feedback was collected and common issues and strengths were identified. The feedback was then used to make modifications to the application. Users considered the updated PortfolioDiet.app as acceptable, giving it a mean SUS score of 85.39 (SD 11.47), which is above the usability quality benchmark threshold score of 70.

In phase 1, the PortfolioDiet.app was found to increase users' perceived knowledge of the Portfolio Diet and to influence their perceived food choices. Responses to open-ended questions revealed common issues and suggestions related to challenges with navigating to (1) resources and (2) profile settings, (3) limited information on plant sterols, (4) inaccuracies in points, (5) timed-logout frustration, (6) request for step-by-step pop-up windows, and (7) request for a mobile app version. When looking at positive feedback, the enjoyment of the recipes was the feature most commonly praised by users. Between the project phases, 6 key modifications were made to the PortfolioDiet.app to incorporate user feedback. In phase 2, the participating users gave the updated PortfolioDiet.app a mean SUS score of 85.39 (SD 11.47). The remaining suggestions to be addressed from phase 1 ([Textbox 1](#): themes 6 and 7) should be prioritized in the next update of the application. Moreover, as identified in the SUS findings in phase 2, the focus should be on engagement and function integration to improve the application's overall usability. To increase engagement, adding social features to the PortfolioDiet.app is recommended. Social features can enhance the benefits of gamification components in engaging users. Patel et al [34] found that a web-based intervention with social support and competition increased physical activity in individuals with type 2 diabetes compared with a control intervention consisting of feedback alone.

Comparison With Previous Work

To our knowledge, this is the first QI initiative undertaken with a nutrition therapy application. There is a paucity of literature focused on QI initiatives with health apps in clinical practice. Although trials investigating the benefits of health apps are common, their findings are inconsistent and the details of their QI initiatives are unclear or not reported. The totality of evidence for web-based applications targeting risk factors for cardiovascular disease found beneficial effects on blood pressure, glycated hemoglobin level, LDL-C, body weight, and physical activity compared with standard of care alone in a systematic review and meta-analysis of 47 randomized controlled trials [15]. However, the evidence for the use of mobile apps to improve health outcomes, although positive, was considered weak based on a recent systematic review and meta-analysis [35]. When looking specifically at mobile apps targeting nutrition-related behaviors, Villinger et al [21] found benefits on both nutrition behaviors and nutrition-related health outcomes; however, these benefits were only found in short-term studies lasting for <6 months. The lack of benefits found in longer-term nutrition app studies may be related to low user engagement because of app usability barriers [22]. Nutrition apps are particularly susceptible to usability issues because they require the user to manually enter food data to provide the user with feedback compared with apps that link to accelerometers and other wearable health devices, such as physical activity apps. These inconsistent findings demonstrate the importance of QI and usability testing of health apps before conducting costly trials.

Previous studies have assessed the usability of digital dietary assessment tools, but these tools were only intended for dietary intake assessment and not for delivering nutrition therapies [36,37]. Usability testing of other lifestyle therapies has been conducted, including a web-based exercise program for older adults (mean SUS score of 84.2, SD 13.3) [38] and a comparison of 2 web-based interventions to increase physical activity, with mean SUS scores of 61.7 (SD 10.8) and 62.5 (SD 11.1) [39]. Another study assessed usability testing of a lifestyle intervention app in patients with type 1 or type 2 diabetes and found a mean SUS score of 62.0 (SD 18.0) [40].

Strengths and Limitations

The purpose of this QI project was to conduct initial testing of the PortfolioDiet.app and to integrate a diverse group of end users in the development and testing of the application. By collecting both qualitative and quantitative data, this project allowed for a more comprehensive collection of data on the views of, and feedback from, end users. The data collected through this QI project identified several important issues with

the previous application version that were able to be addressed and also provided direction for future development. The expansion of the tailored recipes may be especially important because of concerns in the literature regarding the nutritional content of internet recipes [41]. The COVID-19 pandemic has shown that maintaining care at a distance was not only essential, but must also be done well. The resulting application would strongly support distance care both in times of the pandemic and beyond.

A limitation of our study was that we used a convenience sample of users. Although convenience samples have been previously found to increase the risk of bias to favor the intervention, we attempted to limit the bias by purposefully reaching out to a broad range of users (patients with hyperlipidemia, family physicians and registered dietitians, the general public, and medical and nutrition students), resulting in an assorted sample. This assorted sample of users may have allowed for more barriers to be identified. The need for more information on plant sterols may have not emerged with a sample of informed patients and staff from a specialist lipid clinic.

Another limitation is that the SUS was not specifically designed to evaluate therapeutic health apps and is recommended to be combined with other usability metrics. Although there are various methods available to test the usability of therapeutic apps, the SUS is commonly used in the literature [42-44], which allows for comparisons with other therapeutic lifestyle intervention apps. In addition, the SUS uses an intuitive 100-point scale for the score, allowing findings to be easily communicated to those outside of the usability field. Another benefit of the SUS is that it can be completed by users in a short period of time with 10 questions. Other questionnaires developed to assess health apps are longer, increasing user response burden. Future assessments of the mobile version of the PortfolioDiet.app will include other questionnaires more specific to mobile health apps, such as the user version of the Mobile Application Rating Scale questionnaire [45].

Conclusions

The continued consultation with knowledge users throughout the development and testing process of the PortfolioDiet.app aligns with participatory research or iKT approaches [17]. The result of this QI project is a clinical tool that better meets the needs of end users. Although the therapeutic benefits of the Portfolio Diet are well established and the PortfolioDiet.app was demonstrated to increase knowledge of the Portfolio Diet and is usable, the impact of the PortfolioDiet.app on LDL-C and cardiovascular risk reduction is unknown. Therefore, the next step will be to evaluate the utility of the PortfolioDiet.app in primary care settings in a clinical trial.

Acknowledgments

The authors would like to thank all participating users for their time and detailed feedback. The authors would also like to thank their PortfolioDiet.app team and the many volunteers who assisted in the development and formatting of recipes and other important resources used in the application [46]. The Toronto 3D Knowledge Synthesis and Clinical Trials foundation provided funding for the project. The Diet, Digestive tract, and Disease (3D) Centre, funded through the Canada Foundation for Innovation and the Ministry of Research and Innovation's Ontario Research Fund, provided the infrastructure for the conduct of this work.

MEK was funded by a Toronto 3D Knowledge Synthesis and Clinical Trials foundation PhD Scholarship Award. LC was funded by a Mitacs-Elevate Postdoctoral Fellowship Award. AJG was supported by a Canadian Institute of Health Research Fellowship. GH was funded by a Toronto 3D Summer Student Scholarship. DJAJ was funded by the Government of Canada through the Canada Research Chair Endowment. JLS was funded by a PSI Graham Farquharson Knowledge Translation Fellowship; Canadian Diabetes Association Clinician Scientist Award; Canadian Institute of Health Research INMD and CNS New Investigator Partnership Prize; and Banting & Best Diabetes Centre Sun Life Financial New Investigator Award.

Authors' Contributions

All authors were involved in the development and testing of the PortfolioDiet.app prototypes. MEK, LC, AJG, GH, CMC, and JLS met weekly to discuss feedback and prototype updates. MEK contributed to the project design and questionnaire development, analyzed and interpreted the data, wrote the first draft of the manuscript, incorporated revisions, and approved the final version. LC acquired the data, contributed to the project design and questionnaire development, analyzed and interpreted the data, revised the manuscript for important intellectual content, and approved the final version. AJG contributed to the project design, analyzed and interpreted the data, revised the manuscript for important intellectual content, and approved the final version. GH assisted with acquiring the data, contributed to the project design, analyzed and interpreted the data, revised the manuscript for important intellectual content, and approved the final version. CMC developed the different prototypes, revised the manuscript for important intellectual content, and approved the final version. SMG led questionnaire and methods development, revised the manuscript for important intellectual content, and approved the final version. WW, AL, CH, JR, and KS were part of the physician advisory committee and were involved in the design of the application. DS provided expertise in behavior change theory. RGJ, VSM, WW, AL, CH, JR, KS, DS, ESC, JAU, PJ, GLB, MEF, LAL, CWCK, and DJAJ provided input into the project and application design, revised the manuscript for important intellectual content, and approved the final version. JLS provided project supervision, revised the manuscript for important intellectual content, and approved the final version.

Conflicts of Interest

MEK has received funding from a Toronto 3D Knowledge Synthesis and Clinical Trials foundation PhD scholarship award and is a part-time employee at Inquis Clinical Research, a contract research organization. LC was a Mitacs Elevate postdoctoral fellow jointly funded by the government of Canada and Canadian Sugar Institute (from February 2019 to August 2021). AJG has received consulting fees from SoLo GI Nutrition and an honorarium from the Soy Nutrition Institute. PJ serves as an unpaid member of the steering group of trials funded by Appili Therapeutics (Control COVID-19 trial: ClinicalTrials.gov NCT04448119), Abbott Vascular (Evaluation of XIENCE Versus Coronary Artery Bypass Surgery for Effectiveness of Left Main Revascularization [EXCEL] trial: ClinicalTrials.gov NCT01205776; comparing XIENCE stent in participants with unprotected left main coronary artery disease with coronary artery bypass graft surgery; no active involvement for >3 years, no coauthored publication, but still listed as an original member of the statistical executive committee), and Terumo (Management of High Bleeding Risk Patients Post Bioresorbable Polymer Coated Stent Implantation With an Abbreviated Versus Standard DAPT [dual antiplatelet therapy] Regimen [MASTER DAPT] trial: ClinicalTrials.gov NCT03023020; comparing abbreviated DAPT with prolonged DAPT in patients with a drug-eluting stent; ongoing active involvement as a member of the steering group). MEF has received research support from Amgen; Astra Zeneca; Novartis; and Novo Nordisk. CWCK has received grants or research support from the Advanced Foods and Materials Network, Agriculture and Agri-Food Canada, Almond Board of California, American Pistachio Growers, Barilla, Calorie Control Council, Canadian Institutes of Health Research, Canola Council of Canada, International Nut and Dried Fruit Council, International Tree Nut Council Nutrition Research & Education Foundation, Loblaw Companies Ltd, Pulse Canada, Saskatchewan Pulse Growers Association, and Unilever. He has received in-kind research support from the Almond Board of California, American Peanut Council, Barilla, California Walnut Commission, Kellogg Canada, Loblaw Companies Ltd, Quaker (PepsiCo), Primo, Unico, Unilever, and WhiteWave Foods; has received travel support or honoraria from the American Peanut Council, American Pistachio Growers, Barilla, California Walnut Commission, Canola Council of Canada, General Mills, International Nut and Dried Fruit Council, International Pasta Organization, Loblaw Companies Ltd, Nutrition Foundation of Italy, Oldways Preservation Trust, Paramount Farms, Peanut Institute, Pulse Canada, Sabra Dipping, Saskatchewan Pulse Growers Association, Sun-Maid, Tate & Lyle, Unilever, and WhiteWave Foods. He has served on the scientific advisory board for the International Tree Nut Council, International Pasta Organization, Lantmannen, McCormick Science Institute, Oldways Preservation Trust, Paramount Farms, and Pulse Canada. He is a member of the International Carbohydrate Quality Consortium, executive board member of the diabetes and nutrition study group of the European Association for the Study of Diabetes; is on the clinical practice guidelines expert committee for nutrition therapy of the European Association for the Study of Diabetes; and is a director of the Toronto 3D Knowledge Synthesis and Clinical Trials foundation. DJAJ has received research grants from the Saskatchewan and Alberta Pulse Growers Associations; the Agricultural Bioproducts Innovation Program through the Pulse Research Network; Advanced Foods and Materials Network; Loblaw Companies Ltd; Unilever Canada and Netherlands; Barilla; Almond Board of California; Agriculture and Agri-Food Canada; Pulse Canada; Kellogg Canada; Quaker Oats Canada; Procter & Gamble Technical Centre Ltd; Bayer Consumer Care, Springfield, New Jersey; Quaker (PepsiCo); International Nut and Dried Fruit Council; Soy Foods Association of North America; the Coca-Cola Company (investigator-initiated, unrestricted grant); Solae; Haine Celestial; the Sanitarium Company; Orafit; International Tree Nut Council Nutrition Research & Education Foundation; Peanut Institute;

Soy Nutrition Institute; Canola and Flax Councils of Canada; Calorie Control Council; Canadian Institutes of Health Research; Canada Foundation for Innovation; and the Ontario Research Fund. He has received in-kind supplies for trials as research support from the Almond Board of California, Walnut Council of California, Peanut Institute, Barilla, Unilever, Unico, Primo, Loblaw Companies Ltd, Quaker (PepsiCo), Pristine Gourmet, Bunge Limited, Kellogg Canada, and WhiteWave Foods. He has been on the speakers' panel, served on the scientific advisory board, and received travel support and honoraria from Nutritional Fundamentals for Health–Nutramedica; Saint Barnabas Medical Center; the University of Chicago; 2020 China Glycemic Index International Conference; Atlantic Pain Conference; Academy of Life Long Learning; Almond Board of California; Canadian Agriculture Policy Institute; Loblaw Companies Ltd; Griffin Hospital (for the development of the NuVal scoring system); the Coca-Cola Company; Epicure; Danone; Diet Quality Photo Navigation; Better Therapeutics (formerly known as FareWell); Verywell; True Health Initiative; Heali AI Corp; Institute of Food Technologists; Soy Nutrition Institute; Herbalife Nutrition Institute; the Saskatchewan and Alberta Pulse Growers Associations; Sanitarium Company; Orafiti; International Tree Nut Council Nutrition Research & Education Foundation; Peanut Institute; Herbalife International; Pacific Health Laboratories; Barilla; Metagenics; Bayer Consumer Care; Unilever Canada and Netherlands; Solae; Kellogg; Quaker Oats; Procter & Gamble; Abbott Laboratories; Dean Foods; California Strawberry Commission; Haine Celestial; PepsiCo; Alpro Foundation; Pioneer Hi-Bred International; DuPont Nutrition and Health; Spherix Consulting; WhiteWave Foods; Advanced Foods and Materials Network; Canola and Flax Councils of Canada; Agriculture and Agri-Food Canada; Canadian Agri-Food Policy Institute; Pulse Canada; Soy Foods Association of North America; Nutrition Foundation of Italy; Nutra-Source Diagnostics; the McDougall Program; Toronto Knowledge Translation Group (St Michael's Hospital); Canadian College of Naturopathic Medicine; the Hospital for Sick Children; Canadian Nutrition Society; American Society for Nutrition; Arizona State University; Paolo Sorbini Foundation; and Institute of Nutrition, Metabolism and Diabetes. He received an honorarium from the US Department of Agriculture to present the 2013 WO Atwater Memorial Lecture. He received the 2013 Award for Excellence in Research from the International Nut and Dried Fruit Council. He received funding and travel support from the Canadian Society of Endocrinology and Metabolism to produce mini-cases for the Canadian Diabetes Association. He is a member of the International Carbohydrate Quality Consortium. His wife, Alexandra L Jenkins, is a director and partner of Inquis Clinical Research for the food industry; his two daughters, Wendy Jenkins and Amy Jenkins, have published a book for vegetarians that promotes the use of the foods described here, *The Portfolio Diet for Cardiovascular Risk Reduction* (Academic Press [Elsevier] 2020, ISBN: 978-0-12-810510-8); and his sister, Caroline Brydson, received funding through a grant from the St Michael's Hospital Foundation to develop a cookbook for one of his studies. He is also a vegan. JLS has received research support from the Canadian Foundation for Innovation; Ontario Research Fund; Province of Ontario Ministry of Research, Innovation and Science; Canadian Institutes of Health Research; Diabetes Canada; PSI Foundation; Banting & Best Diabetes Centre; American Society for Nutrition; International Nut and Dried Fruit Council Foundation; National Dried Fruit Trade Association; National Honey Board (the US Department of Agriculture honey Checkoff program); International Life Sciences Institute; Pulse Canada; Quaker Oats Center of Excellence; United Soybean Board (the US Department of Agriculture soy Checkoff program); the Tate & Lyle Nutritional Research Fund at the University of Toronto; the Glycemic Control and Cardiovascular Disease in Type 2 Diabetes Fund at the University of Toronto (a fund established by the Alberta Pulse Growers Association); and the Nutrition Trialists Fund at the University of Toronto (a fund established by an inaugural donation from the Calorie Control Council). He has received in-kind food donations to support a randomized controlled trial from the Almond Board of California, California Walnut Commission, Peanut Institute, Barilla, Upfield, Unilever, Unico, Primo, Loblaw Companies Ltd, Quaker, Kellogg Canada, WhiteWave Foods (Danone), and Nutrartis. He has received travel support, speaker fees, and honoraria from Diabetes Canada, Dairy Farmers of Canada, FoodMinds LLC, International Sweeteners Association, Nestlé, Pulse Canada, Canadian Society for Endocrinology and Metabolism, GI Foundation, Abbott, General Mills, Biofortis, the American Society for Nutrition, Northern Ontario School of Medicine, International Tree Nut Council Nutrition Research & Education Foundation, European Food Safety Authority, Comité Européen des Fabricants de Sucre, Nutrition Communications, International Food Information Council, Calorie Control Council, and Physicians Committee for Responsible Medicine. He has, or has had, ad hoc consulting arrangements with Perkins Coie LLP, Tate & Lyle, Wirtschaftliche Vereinigung Zucker eV, Danone, and Inquis Clinical Research. He is a member of the European Fruit Juice Association Scientific Expert Panel and a former member of the Soy Nutrition Institute scientific advisory committee. He is on the clinical practice guidelines expert committees of Diabetes Canada, European Association for the Study of Diabetes, and Canadian Cardiovascular Society, as well as Obesity Canada and the Canadian Association of Bariatric Physicians and Surgeons. He serves, or has served, as an unpaid scientific advisor for the food, nutrition, and safety program and the technical committee on carbohydrates of International Life Sciences Institute North America. He is a member of the International Carbohydrate Quality Consortium, executive board member of the diabetes and nutrition study group of the European Association for the Study of Diabetes, and director of the Toronto 3D Knowledge Synthesis and Clinical Trials foundation. His wife is an employee of AB InBev. Authors GH, CMC, SMG, RGJ, VSM, WW, AL, CH, JR, KS, DS, ESC, JAU, GLB, LAL have no conflicts of interest to disclose.

Multimedia Appendix 1

Screenshots depicting the various features on the dashboard of the application.

[[PDF File \(Adobe PDF File\), 760 KB](#) - [humanfactors_v9i2e34704_app1.pdf](#)]

Multimedia Appendix 2

The Portfolio Diet infographic.

[[PDF File \(Adobe PDF File\), 2586 KB - humanfactors_v9i2e34704_app2.pdf](#)]

Multimedia Appendix 3

Example pages from the instructional guide (navigation and progress).

[[PDF File \(Adobe PDF File\), 363 KB - humanfactors_v9i2e34704_app3.pdf](#)]

Multimedia Appendix 4

Mixed-form feedback questionnaire (phase 1).

[[DOCX File , 20 KB - humanfactors_v9i2e34704_app4.docx](#)]

Multimedia Appendix 5

System Usability Scale (phase 2).

[[DOCX File , 19 KB - humanfactors_v9i2e34704_app5.docx](#)]

Multimedia Appendix 6

Tables showing full quantitative responses (phase 1), full qualitative responses broken up by limitations and strengths (phase 1), and scores for individual System Usability Scale items (phase 2).

[[DOCX File , 36 KB - humanfactors_v9i2e34704_app6.docx](#)]

Multimedia Appendix 7

Screenshots depicting updates made to the application based on user feedback.

[[PDF File \(Adobe PDF File\), 2420 KB - humanfactors_v9i2e34704_app7.pdf](#)]

References

1. Jenkins DJ, Kendall CW, Marchie A, Faulkner DA, Wong JM, de Souza R, et al. Effects of a dietary portfolio of cholesterol-lowering foods vs lovastatin on serum lipids and C-reactive protein. *JAMA* 2003;290(4):502-510. [doi: [10.1001/jama.290.4.502](#)] [Medline: [12876093](#)]
2. Jenkins DJ, Jones PJ, Lamarche B, Kendall CW, Faulkner D, Cermakova L, et al. Effect of a dietary portfolio of cholesterol-lowering foods given at 2 levels of intensity of dietary advice on serum lipids in hyperlipidemia: a randomized controlled trial. *JAMA* 2011;306(8):831-839. [doi: [10.1001/jama.2011.1202](#)] [Medline: [21862744](#)]
3. Chiavaroli L, Nishi SK, Khan TA, Braunstein CR, Glenn AJ, Mejia SB, et al. Portfolio dietary pattern and cardiovascular disease: a systematic review and meta-analysis of controlled trials. *Prog Cardiovasc Dis* 2018;61(1):43-53 [[FREE Full text](#)] [doi: [10.1016/j.pcad.2018.05.004](#)] [Medline: [29807048](#)]
4. Pearson G, Thanassoulis G, Anderson T, Barry AR, Couture P, Dayan N, et al. 2021 Canadian cardiovascular society guidelines for the management of dyslipidemia for the prevention of cardiovascular disease in adults. *Can J Cardiol* 2021;37(8):1129-1150. [doi: [10.1016/j.cjca.2021.03.016](#)] [Medline: [33781847](#)]
5. Anderson TJ, Grégoire J, Pearson GJ, Barry AR, Couture P, Dawes M, et al. 2016 Canadian cardiovascular society guidelines for the management of dyslipidemia for the prevention of cardiovascular disease in the adult. *Can J Cardiol* 2016;32(11):1263-1282. [doi: [10.1016/j.cjca.2016.07.510](#)] [Medline: [27712954](#)]
6. Diabetes Canada Clinical Practice Guidelines Expert Committee, Punthakee Z, Goldenberg R, Katz P. Definition, classification and diagnosis of diabetes, prediabetes and metabolic syndrome. *Can J Diabetes* 2018;42 Suppl 1:S10-S15. [doi: [10.1016/j.cjcd.2017.10.003](#)] [Medline: [29650080](#)]
7. Brown J, Clarke C, Johnson Stoklossa V, Sievenpiper J. Canadian adult obesity clinical practice guidelines: medical nutrition therapy in obesity management. Obesity Canada. 2020. URL: <https://obesitycanada.ca/guidelines/nutrition> [accessed 2021-03-25]
8. Tobe SW, Stone JA, Anderson T, Bacon S, Cheng AY, Daskalopoulou SS, et al. Canadian cardiovascular harmonized national guidelines endeavour (C-CHANGE) guideline for the prevention and management of cardiovascular disease in primary care: 2018 update. *CMAJ* 2018;190(40):E1192-E1206 [[FREE Full text](#)] [doi: [10.1503/cmaj.180194](#)] [Medline: [30301743](#)]
9. Portfolio diet. Heart UK: The cholesterol charity. URL: <https://www.heartuk.org.uk/downloads/healthprofessionals/factsheets/the-portfolio-diet.pdf> [accessed 2016-02-04]
10. Stroes ES, Thompson PD, Corsini A, Vladutiu GD, Raal FJ, Ray KK, European Atherosclerosis Society Consensus Panel. Statin-associated muscle symptoms: impact on statin therapy-European atherosclerosis society consensus panel statement on assessment, aetiology and management. *Eur Heart J* 2015;36(17):1012-1022 [[FREE Full text](#)] [doi: [10.1093/eurheartj/ehv043](#)] [Medline: [25694464](#)]

11. Grundy SM, Stone NJ, Bailey AL, Beam C, Birtcher KK, Blumenthal RS, et al. 2018 AHA/ACC/AACVPR/AAPA/ABC/ACPM/ADA/AGS/APhA/ASPC/NLA/PCNA guideline on the management of blood cholesterol: executive summary: a report of the American College of Cardiology/American Heart Association Task Force on Clinical Practice Guidelines. *J Am Coll Cardiol* 2019;73(24):3168-3209. [doi: [10.1016/j.jacc.2018.11.002](https://doi.org/10.1016/j.jacc.2018.11.002)] [Medline: [30423391](https://pubmed.ncbi.nlm.nih.gov/30423391/)]
12. Kris-Etherton PM, Akabas SR, Bales CW, Bistrian B, Braun L, Edwards MS, et al. The need to advance nutrition education in the training of health care professionals and recommended research to evaluate implementation and effectiveness. *Am J Clin Nutr* 2014;99(5 Suppl):1153S-1166S [FREE Full text] [doi: [10.3945/ajcn.113.073502](https://doi.org/10.3945/ajcn.113.073502)] [Medline: [24717343](https://pubmed.ncbi.nlm.nih.gov/24717343/)]
13. Aboueid S, Bourgeault I, Giroux I. Nutrition and obesity care in multidisciplinary primary care settings in Ontario, Canada: short duration of visits and complex health problems perceived as barriers. *Prev Med Rep* 2018;10:242-247 [FREE Full text] [doi: [10.1016/j.pmedr.2018.04.003](https://doi.org/10.1016/j.pmedr.2018.04.003)] [Medline: [29868375](https://pubmed.ncbi.nlm.nih.gov/29868375/)]
14. Block G, Azar KM, Romanelli RJ, Block TJ, Hopkins D, Carpenter HA, et al. Diabetes prevention and weight loss with a fully automated behavioral intervention by email, web, and mobile phone: a randomized controlled trial among persons with prediabetes. *J Med Internet Res* 2015;17(10):e240 [FREE Full text] [doi: [10.2196/jmir.4897](https://doi.org/10.2196/jmir.4897)] [Medline: [26499966](https://pubmed.ncbi.nlm.nih.gov/26499966/)]
15. Beishuizen CR, Stephan BC, van Gool WA, Brayne C, Peters RJ, Andrieu S, et al. Web-based interventions targeting cardiovascular risk factors in middle-aged and older people: a systematic review and meta-analysis. *J Med Internet Res* 2016;18(3):e55 [FREE Full text] [doi: [10.2196/jmir.5218](https://doi.org/10.2196/jmir.5218)] [Medline: [26968879](https://pubmed.ncbi.nlm.nih.gov/26968879/)]
16. Bruining N. The post-pandemic legacy: the breakthrough of digital health and telemedicine. *Cardiovasc Res* 2021;117(9):e118-e119 [FREE Full text] [doi: [10.1093/cvr/cvab178](https://doi.org/10.1093/cvr/cvab178)] [Medline: [34143873](https://pubmed.ncbi.nlm.nih.gov/34143873/)]
17. Gagliardi AR, Berta W, Kothari A, Boyko J, Urquhart R. Integrated knowledge translation (IKT) in health care: a scoping review. *Implement Sci* 2016;11(1):38 [FREE Full text] [doi: [10.1186/s13012-016-0399-1](https://doi.org/10.1186/s13012-016-0399-1)] [Medline: [26988000](https://pubmed.ncbi.nlm.nih.gov/26988000/)]
18. Quality improvement guide. Health Quality Ontario (HQO). 2012. URL: <https://www.hqontario.ca/portals/0/documents/qi/qi-quality-improve-guide-2012-en.pdf> [accessed 2021-09-13]
19. Brown 3rd W, Yen PY, Rojas M, Schnall R. Assessment of the health IT usability evaluation model (Health-ITUEM) for evaluating mobile health (mHealth) technology. *J Biomed Inform* 2013;46(6):1080-1087 [FREE Full text] [doi: [10.1016/j.jbi.2013.08.001](https://doi.org/10.1016/j.jbi.2013.08.001)] [Medline: [23973872](https://pubmed.ncbi.nlm.nih.gov/23973872/)]
20. Eysenbach G. The law of attrition. *J Med Internet Res* 2005;7(1):e11 [FREE Full text] [doi: [10.2196/jmir.7.1.e11](https://doi.org/10.2196/jmir.7.1.e11)] [Medline: [15829473](https://pubmed.ncbi.nlm.nih.gov/15829473/)]
21. Villinger K, Wahl DR, Boeing H, Schupp HT, Renner B. The effectiveness of app-based mobile interventions on nutrition behaviours and nutrition-related health outcomes: a systematic review and meta-analysis. *Obes Rev* 2019;20(10):1465-1484 [FREE Full text] [doi: [10.1111/obr.12903](https://doi.org/10.1111/obr.12903)] [Medline: [31353783](https://pubmed.ncbi.nlm.nih.gov/31353783/)]
22. König LM, Attig C, Franke T, Renner B. Barriers to and facilitators for using nutrition apps: systematic review and conceptual framework. *JMIR Mhealth Uhealth* 2021;9(6):e20037 [FREE Full text] [doi: [10.2196/20037](https://doi.org/10.2196/20037)] [Medline: [34254938](https://pubmed.ncbi.nlm.nih.gov/34254938/)]
23. The Portfolio Diet. 2021. URL: <https://portfoliodiet.app/> [accessed 2021-11-03]
24. Paré G, Bourget C. Diffusion of smart devices for health in Canada. *Canada Health Infoway*. 2017. URL: <https://tinyurl.com/bdd3n3en> [accessed 2021-10-07]
25. Coorey GM, Neubeck L, Mulley J, Redfern J. Effectiveness, acceptability and usefulness of mobile applications for cardiovascular disease self-management: systematic review with meta-synthesis of quantitative and qualitative data. *Eur J Prev Cardiol* 2018;25(5):505-521. [doi: [10.1177/2047487317750913](https://doi.org/10.1177/2047487317750913)] [Medline: [29313363](https://pubmed.ncbi.nlm.nih.gov/29313363/)]
26. Curry LA, Krumholz HM, O' Cathain A, Plano Clark VL, Cherlin E, Bradley EH. Mixed methods in biomedical and health services research. *Circ Cardiovasc Qual Outcomes* 2013;6(1):119-123. [doi: [10.1161/CIRCOUTCOMES.112.967885](https://doi.org/10.1161/CIRCOUTCOMES.112.967885)] [Medline: [23322807](https://pubmed.ncbi.nlm.nih.gov/23322807/)]
27. Brooke J. SUS: a 'quick and dirty' usability scale. In: Jordan PW, Thomas B, McClelland I, Weerdmeester B, editors. *Usability evaluation in industry*. Boca Raton, FL: CRC Press; 1996.
28. Lewis JR, Sauro J. The factor structure of the system usability scale. In: *International Conference on Human Centered Design*. 2009 Presented at: HCD '09; July 19-24, 2009; San Diego, CA p. 94-103. [doi: [10.1007/978-3-642-02806-9_12](https://doi.org/10.1007/978-3-642-02806-9_12)]
29. Bangor A, Kortum PT, Miller JT. An empirical evaluation of the system usability scale. *Int J Hum Comput Interact* 2008;24(6):574-594. [doi: [10.1080/10447310802205776](https://doi.org/10.1080/10447310802205776)]
30. Nielsen J. Why you only need to test with 5 users. Nielsen Norman Group. 2000. URL: <https://www.nngroup.com/articles/how-many-test-users/> [accessed 2021-10-13]
31. Macefield R. How to specify the participant group size for usability studies: a practitioner's guide. *J Usability Stud* 2009;5(1):34-45.
32. Melin J, Bonn SE, Pendrill L, Trolle Lagerros Y. A questionnaire for assessing user satisfaction with mobile health apps: development using Rasch measurement theory. *JMIR Mhealth Uhealth* 2020;8(5):e15909 [FREE Full text] [doi: [10.2196/15909](https://doi.org/10.2196/15909)] [Medline: [32452817](https://pubmed.ncbi.nlm.nih.gov/32452817/)]
33. Graneheim UH, Lundman B. Qualitative content analysis in nursing research: concepts, procedures and measures to achieve trustworthiness. *Nurse Educ Today* 2004;24(2):105-112. [doi: [10.1016/j.nedt.2003.10.001](https://doi.org/10.1016/j.nedt.2003.10.001)] [Medline: [14769454](https://pubmed.ncbi.nlm.nih.gov/14769454/)]

34. Patel MS, Small DS, Harrison JD, Hilbert V, Fortunato MP, Oon AL, et al. Effect of behaviorally designed gamification with social incentives on lifestyle modification among adults with uncontrolled diabetes: a randomized clinical trial. *JAMA Netw Open* 2021;4(5):e2110255 [FREE Full text] [doi: [10.1001/jamanetworkopen.2021.10255](https://doi.org/10.1001/jamanetworkopen.2021.10255)] [Medline: [34028550](https://pubmed.ncbi.nlm.nih.gov/34028550/)]
35. Iribarren SJ, Akande TO, Kamp KJ, Barry D, Kader YG, Suelzer E. Effectiveness of mobile apps to promote health and manage disease: systematic review and meta-analysis of randomized controlled trials. *JMIR Mhealth Uhealth* 2021;9(1):e21563 [FREE Full text] [doi: [10.2196/21563](https://doi.org/10.2196/21563)] [Medline: [33427672](https://pubmed.ncbi.nlm.nih.gov/33427672/)]
36. Albar SA, Carter MC, Alwan NA, Evans CE, Cade JE. Formative evaluation of the usability and acceptability of myfood24 among adolescents: a UK online dietary assessments tool. *BMC Nutr* 2015;1(1):1. [doi: [10.1186/s40795-015-0016-8](https://doi.org/10.1186/s40795-015-0016-8)]
37. Zenun Franco R, Fallaize R, Lovegrove JA, Hwang F. Online dietary intake assessment using a graphical food frequency app (eNutri): usability metrics from the EatWellUK study. *PLoS One* 2018;13(8):e0202006 [FREE Full text] [doi: [10.1371/journal.pone.0202006](https://doi.org/10.1371/journal.pone.0202006)] [Medline: [30096211](https://pubmed.ncbi.nlm.nih.gov/30096211/)]
38. Dekker-van Weering M, Jansen-Kosterink S, Frazer S, Vollenbroek-Hutten M. User experience, actual use, and effectiveness of an information communication technology-supported home exercise program for pre-frail older adults. *Front Med (Lausanne)* 2017;4:208 [FREE Full text] [doi: [10.3389/fmed.2017.00208](https://doi.org/10.3389/fmed.2017.00208)] [Medline: [29250523](https://pubmed.ncbi.nlm.nih.gov/29250523/)]
39. Vandelanotte C, Kolt GS, Caperchione CM, Savage TN, Rosenkranz RR, Maeder AJ, et al. Effectiveness of a web 2.0 intervention to increase physical activity in real-world settings: randomized ecological trial. *J Med Internet Res* 2017;19(11):e390 [FREE Full text] [doi: [10.2196/jmir.8484](https://doi.org/10.2196/jmir.8484)] [Medline: [29133282](https://pubmed.ncbi.nlm.nih.gov/29133282/)]
40. Fu HN, Adam TJ, Konstan JA, Wolfson JA, Clancy TR, Wyman JF. Influence of patient characteristics and psychological needs on diabetes mobile app usability in adults with type 1 or type 2 diabetes: crossover randomized trial. *JMIR Diabetes* 2019;4(2):e11462 [FREE Full text] [doi: [10.2196/11462](https://doi.org/10.2196/11462)] [Medline: [31038468](https://pubmed.ncbi.nlm.nih.gov/31038468/)]
41. Trattner C, Elswailer D, Howard S. Estimating the healthiness of internet recipes: a cross-sectional study. *Front Public Health* 2017;5:16 [FREE Full text] [doi: [10.3389/fpubh.2017.00016](https://doi.org/10.3389/fpubh.2017.00016)] [Medline: [28243587](https://pubmed.ncbi.nlm.nih.gov/28243587/)]
42. Edney SM, Olds TS, Ryan JC, Vandelanotte C, Plotnikoff RC, Curtis RG, et al. A social networking and gamified app to increase physical activity: cluster RCT. *Am J Prev Med* 2020;58(2):e51-e62. [doi: [10.1016/j.amepre.2019.09.009](https://doi.org/10.1016/j.amepre.2019.09.009)] [Medline: [31959326](https://pubmed.ncbi.nlm.nih.gov/31959326/)]
43. Ormel HL, van der Schoot GG, Westerink NL, Sluiter WJ, Gietema JA, Walenkamp AM. Self-monitoring physical activity with a smartphone application in cancer patients: a randomized feasibility study (SMART-trial). *Support Care Cancer* 2018;26(11):3915-3923 [FREE Full text] [doi: [10.1007/s00520-018-4263-5](https://doi.org/10.1007/s00520-018-4263-5)] [Medline: [29785635](https://pubmed.ncbi.nlm.nih.gov/29785635/)]
44. Rayward AT, Murawski B, Duncan MJ, Holliday EG, Vandelanotte C, Brown WJ, et al. Efficacy of an m-health physical activity and sleep intervention to improve sleep quality in middle-aged adults: the refresh study randomized controlled trial. *Ann Behav Med* 2020;54(7):470-483. [doi: [10.1093/abm/kaz064](https://doi.org/10.1093/abm/kaz064)] [Medline: [31942918](https://pubmed.ncbi.nlm.nih.gov/31942918/)]
45. Stoyanov SR, Hides L, Kavanagh DJ, Wilson H. Development and validation of the user version of the mobile application rating scale (uMARS). *JMIR Mhealth Uhealth* 2016;4(2):e72 [FREE Full text] [doi: [10.2196/mhealth.5849](https://doi.org/10.2196/mhealth.5849)] [Medline: [27287964](https://pubmed.ncbi.nlm.nih.gov/27287964/)]
46. Our team. The Portfolio Diet App. 2021. URL: <https://recipes.portfoliodiet.app/our-team/> [accessed 2021-09-10]

Abbreviations

- iKT:** integrated knowledge translation
LDL-C: low-density lipoprotein cholesterol
QI: quality improvement
SUS: System Usability Scale

Edited by A Kushniruk; submitted 03.12.21; peer-reviewed by G Block, S Jansen-Kosterink; comments to author 29.12.21; revised version received 04.02.22; accepted 13.02.22; published 21.04.22.

Please cite as:

Kavanagh ME, Chiavaroli L, Glenn AJ, Heijmans G, Grant SM, Chow CM, Josse RG, Malik VS, Watson W, Lofters A, Holmes C, Rackal J, Srichaikul K, Sherifali D, Snelgrove-Clarke E, Udell JA, Juni P, Booth GL, Farkouh ME, Leiter LA, Kendall CWC, Jenkins DJA, Sievenpiper JL

A Web-Based Health Application to Translate Nutrition Therapy for Cardiovascular Risk Reduction in Primary Care (PortfolioDiet.app): Quality Improvement and Usability Testing Study

JMIR Hum Factors 2022;9(2):e34704

URL: <https://humanfactors.jmir.org/2022/2/e34704>

doi: [10.2196/34704](https://doi.org/10.2196/34704)

PMID: [35451981](https://pubmed.ncbi.nlm.nih.gov/35451981/)

©Meaghan E Kavanagh, Laura Chiavaroli, Andrea J Glenn, Genevieve Heijmans, Shannan M Grant, Chi-Ming Chow, Robert G Josse, Vasanti S Malik, William Watson, Aisha Lofters, Candice Holmes, Julia Rackal, Kristie Srichaikul, Diana Sherifali, Erna Snelgrove-Clarke, Jacob A Udell, Peter Juni, Gillian L Booth, Michael E Farkouh, Lawrence A Leiter, Cyril WC Kendall, David JA Jenkins, John L Sievenpiper. Originally published in JMIR Human Factors (<https://humanfactors.jmir.org>), 21.04.2022. This is an open-access article distributed under the terms of the Creative Commons Attribution License (<https://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Human Factors, is properly cited. The complete bibliographic information, a link to the original publication on <https://humanfactors.jmir.org>, as well as this copyright and license information must be included.

Original Paper

Effective Communication of Personalized Risks and Patient Preferences During Surgical Informed Consent Using Data Visualization: Qualitative Semistructured Interview Study With Patients After Surgery

Undina Gisladdottir¹, MBI; Drashko Nakikj¹, PhD; Rashi Jhunjunwala², MA, MD; Jasmine Panton^{2,3}, MBE; Gabriel Brat^{1,2*}, MPH, MD; Nils Gehlenborg^{1*}, PhD

¹Department of Biomedical Informatics, Harvard Medical School, Harvard University, Boston, MA, United States

²Department of Surgery, Beth Israel Deaconess Medical Center, Boston, MA, United States

³Geisel School of Medicine, Dartmouth College, Hanover, NH, United States

* these authors contributed equally

Corresponding Author:

Nils Gehlenborg, PhD

Department of Biomedical Informatics

Harvard Medical School

Harvard University

10 Shattuck Street

Suite 514

Boston, MA, 02115

United States

Phone: 1 6174321535

Email: nils@hms.harvard.edu

Abstract

Background: There is no consensus on which risks to communicate to a prospective surgical patient during informed consent or how. Complicating the process, patient preferences may diverge from clinical assumptions and are often not considered for discussion. Such discrepancies can lead to confusion and resentment, raising the potential for legal action. To overcome these issues, we propose a visual consent tool that incorporates patient preferences and communicates personalized risks to patients using data visualization. We used this platform to identify key effective visual elements to communicate personalized surgical risks.

Objective: Our main focus is to understand how to best communicate personalized risks using data visualization. To contextualize patient responses to the main question, we examine how patients perceive risks before surgery (research question 1), how suitably the visual consent tool is able to present personalized surgical risks (research question 2), how well our visualizations convey those personalized surgical risks (research question 3), and how the visual consent tool could improve the informed consent process and how it can be used (research question 4).

Methods: We designed a visual consent tool to meet the objectives of our study. To calculate and list personalized surgical risks, we used the American College of Surgeons risk calculator. We created multiple visualization mock-ups using visual elements previously determined to be well-received for risk communication. Semistructured interviews were conducted with patients after surgery, and each of the mock-ups was presented and evaluated independently and in the context of our visual consent tool design. The interviews were transcribed, and thematic analysis was performed to identify major themes. We also applied a quantitative approach to the analysis to assess the prevalence of different perceptions of the visualizations presented in our tool.

Results: In total, 20 patients were interviewed, with a median age of 59 (range 29-87) years. Thematic analysis revealed *factors that influenced the perception of risk (the surgical procedure, the cognitive capacity of the patient, and the timing of consent;* research question 1); *factors that influenced the perceived value of risk visualizations (preference for rare event communication, preference for risk visualization, and usefulness of comparison with the average;* research question 3); and perceived usefulness and use cases of the visual consent tool (research questions 2 and 4). Most importantly, we found that patients preferred the visual

consent tool to current text-based documents and had no unified preferences for risk visualization. Furthermore, our findings suggest that patient concerns were not often represented in existing risk calculators.

Conclusions: We identified key elements that influence effective visual risk communication in the perioperative setting and pointed out the limitations of the existing calculators in addressing patient concerns. Patient preference is highly variable and should influence choices regarding risk presentation and visualization.

(*JMIR Hum Factors* 2022;9(2):e29118) doi:[10.2196/29118](https://doi.org/10.2196/29118)

KEYWORDS

data visualization; surgical informed consent; shared decision-making; biomedical informatics

Introduction

Background

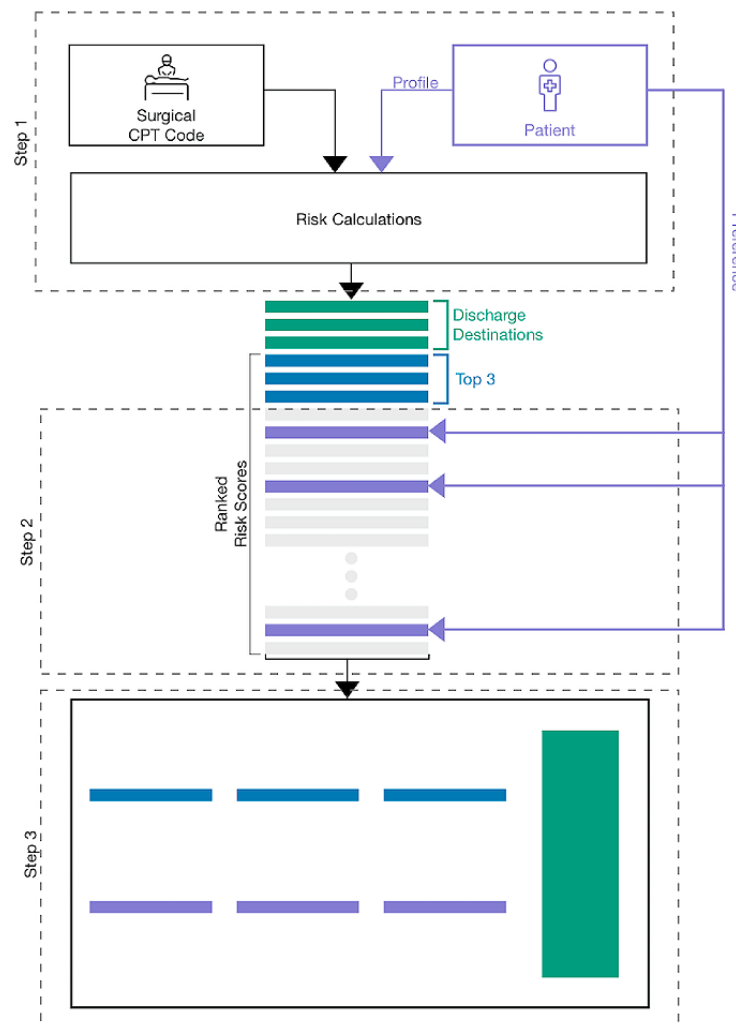
In the United States, >50 million surgical procedures are performed annually [1]. For each procedure, a clinician obtains informed consent from the patient or a surrogate. The discussion during this process plays an important legal and ethical role and should determine the appropriate treatment plan for each patient. The literature suggests that this discussion often does not address the patient's personal treatment goals [2,3]. In addition, many important details are solely communicated verbally [2]. Unexpected, poorly communicated, or possibly life-threatening events can lead to malpractice lawsuits [2,4,5]. Instead, the informed consent conversation should properly set the patient's expectations to decrease the chances of what a patient would consider a nonbeneficial outcome [2,3,5].

Although medical professionals agree that determining patient priorities is important for choosing the appropriate treatment plan, the discussion during informed consent often fails to consider the patient's condition and treatment goals [2,3]. Furthermore, the current informed consent process is not standardized and leaves patients without a clear understanding of the consequences of surgery [6]. There is also a lack of consensus in the medical community regarding which risks to communicate, and risk estimates are often too broad and vary among physicians [2]. Multiple studies have shown that, despite reviewing the surgical procedure and associated risks, the patients' understanding after these discussions is well below acceptable limits [7]. Risk score calculators try to expand the conversation through personalized risks for any given patient. They provide discrete risk scores for a variety of outcomes based on the surgical procedure and preoperative patient data. Despite the growing prevalence of these tools, the surgical community

has not reached a consensus on how to communicate these scores. Some groups have attempted to address this issue by categorizing complications into best case and worst case or good, intermediate, and bad [8-10]. In these approaches, patient preference, which is essential for defining a good outcome for a patient, is not necessarily incorporated or used to inform the conversation.

We propose a design for a visual consent tool to address previous limitations in (1) incorporating patient preferences, (2) setting expectations for the upcoming surgery, and (3) standardizing risk communication during informed consent. The visual consent tool communicates personalized risks to the patients in 3 main steps (Figure 1). First, personalized risks are calculated using one of the risk prediction models currently available [11-14]. These prediction models typically incorporate a surgical Current Procedural Terminology code and patient preoperative data to calculate risks. The design allows for the use of a preferred risk calculator such as the American College of Surgeons (ACS) [13], the Surgical Risk Preoperative Assessment System [12], or the Predictive Optimal Trees in Emergency Surgery Risk [11], among others, without affecting the rest of the workflow. In our particular design instance, we rely on a simulation of the ACS risk calculator at a level that allows us to go through the visual consent tool workflow and conduct our study. Second, patients select a limited number of major concerns (we chose 3 arbitrarily) out of a list of 20 complications produced by the simulated ACS calculator, preranked in descending order of likelihood. Finally, we visualize the probability of the 3 most likely and patient-selected complications as well as the potential discharge destinations: home, rehabilitation, and death. With this, the visual consent tool allows patients to compare the risks of the most likely and prioritized complications and communicates potential discharge destinations after surgery.

Figure 1. The proposed visual consent tool includes 3 main steps to help the patient and surgeon evaluate the risks of a surgery. In the first step, the personalized risks are calculated using an existing risk model with identification of surgery (eg, Current Procedural Terminology [CPT] code) and patient preoperative data as inputs. The patient then chooses up to 3 risks that are of high concern (purple bars) in addition to the top 3 calculated risks (blue bars). Finally, a visualization of these 6 risks is displayed along with the likelihood of each of the final discharge destinations (green bar).



Objectives

Using high-fidelity design mock-ups for the visual consent tool, we conduct a qualitative design feedback study in which we want to address the following research questions: (1) How do patients perceive risks before surgery—does what matters depend on the context? (research question 1), (2) How suitably the visual consent tool is able to present personalized surgical risks—are the patients looking at the risks they really care about the most? (research question 2), (3) How to best communicate these personalized risks using data visualization approaches—are there ways to present these risks that are most understandable to patients? (research question 3), and (4) In which scenarios can the visual consent tool be used, and can it improve the informed consent process—are the patient and surgeon able to engage in a more productive discussion?

Focusing only on the visual consent tool's personalized risk visualization component (Figure 1, step 3), we conduct semistructured interviews with patients during their postoperative visit to the acute care surgical clinic at an

academic medical center. Through thematic analysis of the interviews, we identify several factors that affect the perception of risks and their importance, the perceived value of risk visualization, the preferences for risk visualization, the effects of risk visualization, and the potential usefulness of the visual consent tool in a real-life setting. The report of this study is based on the COREQ (consolidated criteria for reporting qualitative research) guidelines [15].

Methods

Visual Consent Tool Design

A schematic overview of the visual consent tool is shown in Figure 1. The visual consent tool consists of three elements: risk calculation, preference identification, and risk visualization.

Risk Calculation

Multiple methods for calculating personalized perioperative risks for patients have been published [11-14]. These calculators

use collected patient data (eg, age, sex, and smoking status) to calculate the risk of a given postoperative complication.

As an example, the ACS risk calculator, the most commonly used tool, leverages National Surgical Quality Improvement Program participant data from >400 hospitals to calculate 20 different perioperative risks [13]. In this study, we did not focus

on improving risk calculation. Generally, our approach could be applied to risks calculated using any risk calculator. For practical purposes, we used the results obtained from the ACS risk calculator for this study. In our proposed interface, the surgeon provides information about the surgery by entering a Current Procedural Terminology code and the patient profile (Figure 2).

Figure 2. Patient profile and Current Procedural Terminology (CPT) code input. The form at the top is used to enter the CPT code for the surgery, and the form at the bottom is used to provide patient characteristics required by the risk calculator. ASA: American Society of Anesthesia; COPD: chronic obstructive pulmonary disease.

The figure consists of two screenshots of a web application interface. The top screenshot shows a form titled "Data Driven Visual Consent" with navigation links for "Home", "About", and "Team". The main content area prompts the user to "Please enter your procedure:" and features a text input field containing "47563 - Laparoscopic Procedure...". Below the input field is a dark blue "Continue" button. The bottom screenshot shows the same interface after the procedure is entered. It displays "For the following procedure: 47563: Laparoscopic Procedures on the Biliary Tract" and asks the user to "Please fill in the patient information:". The form contains several dropdown menus for "Age Group" (65 years and under), "Sex" (Female), "Body Mass Index (BMI)" (30), "Functional Status" (Independent), "ASA Class" (None Assigned), "Systemic Sepsis" (None), "Diabetes" (None), and "Dyspnea" (No). Below these are 12 light blue buttons for medical conditions: "Emergency", "Steroid use", "Ascites within 30 days", "Ventilator Dependent", "Dialysis", "Disseminated Cancer", "Hypertension w/ medication", "Recent Congestive Heart Failure", "Current Smoker (within 1 year)", "History of Severe COPD", and "Acute Renal Failure". A dark blue "Continue" button is positioned at the bottom center.

Incorporating Patient Preferences

To incorporate personal preferences, our tool provides an interface for patients to identify 3 complications that are of particular concern in addition to the top 3 risks that the tool

automatically selects as most important based on the risk calculations (Figure 3). After presenting the patient with a list of possible complications preranked by likelihood, patients are able to choose the risks that are most concerning to them (Figure 3).

Figure 3. Incorporating patient preferences. A total of 3 most common complications are preselected, with the remaining complications listed in descending order of likelihood. The patient can select up to 3 risks at a time. OR: operating room; SSI: surgical site infection.

Data Driven Visual Consent Home About Team

For the following procedure:
47563

With the following profile:
Male, Ascites, Current Smoker

These are your most likely complications:

Ventilator Return to OR Renal Insufficiency

Choose three complications that would concern you:

Ostomy <input type="checkbox"/>	Stroke <input type="checkbox"/>	Cardiac Event <input type="checkbox"/>
Systemic Sepsis <input type="checkbox"/>	Superficial Infection <input type="checkbox"/>	Organ/Space SSI <input type="checkbox"/>
Unplanned Intubation <input type="checkbox"/>	Acute Renal Failure <input type="checkbox"/>	Pulmonary Embolism <input type="checkbox"/>
Myocardial Infarction <input type="checkbox"/>	Wound Disruption <input type="checkbox"/>	Deep Incisional Infection <input type="checkbox"/>

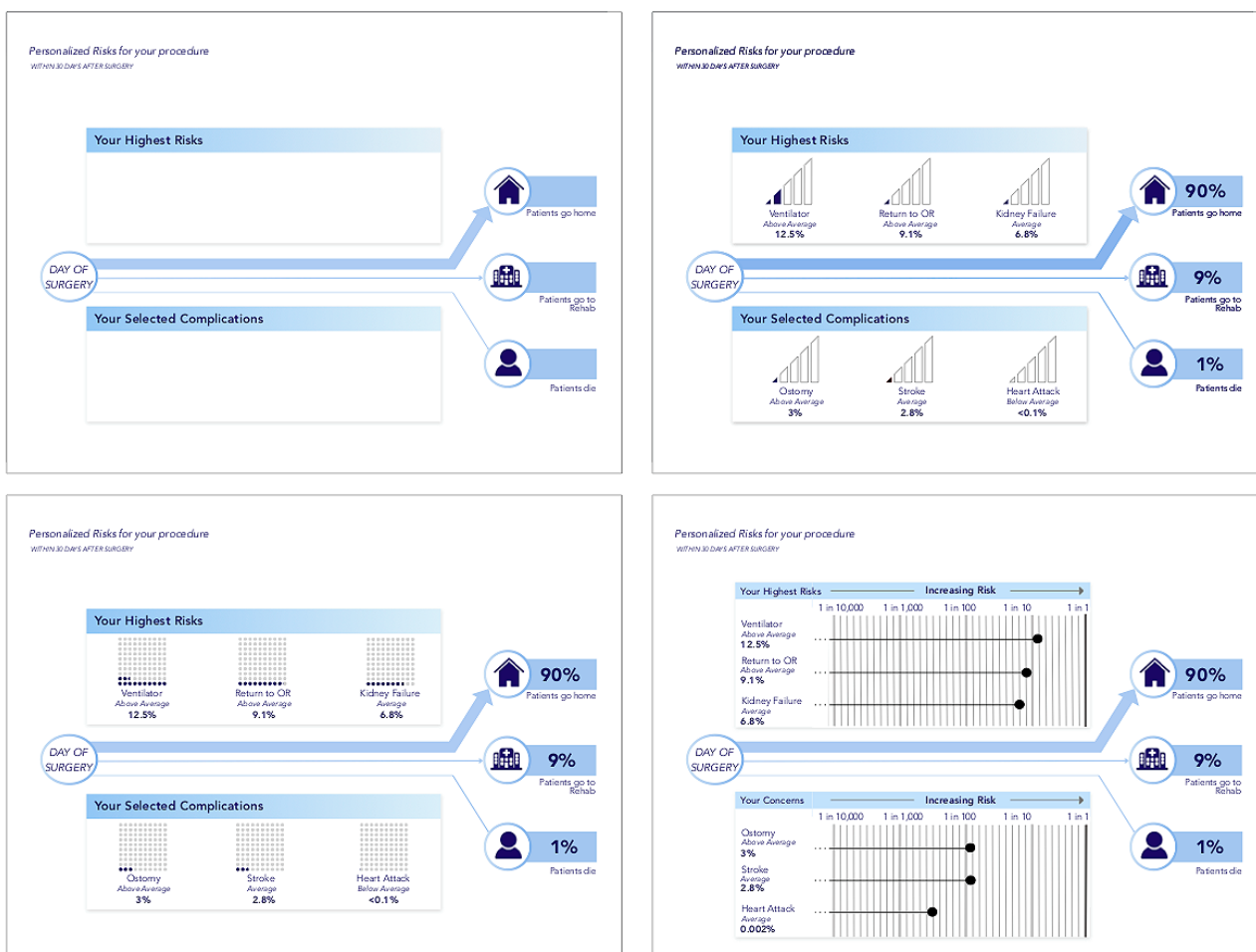
Back Continue

Risk Visualization

The visualization is intended to communicate personalized perioperative risks and the likelihood of the discharge destinations in a clear and understandable manner. The overall goal is to promote a more coherent discussion between the surgeon and patient for improved shared decision-making. The layout includes the most likely preselected complications based on the risk calculations as well as those selected by the patient

and the likelihood of each discharge destination (Figure 4, top left). Discharge destinations are communicated using weighted lines to represent likelihood. Preselected and patient-selected complications are boxed separately to allow for comparison between the 2 categories. Given the relatively low rates of complications, the representation of the likelihood of each complication presented a unique challenge that we examined in detail.

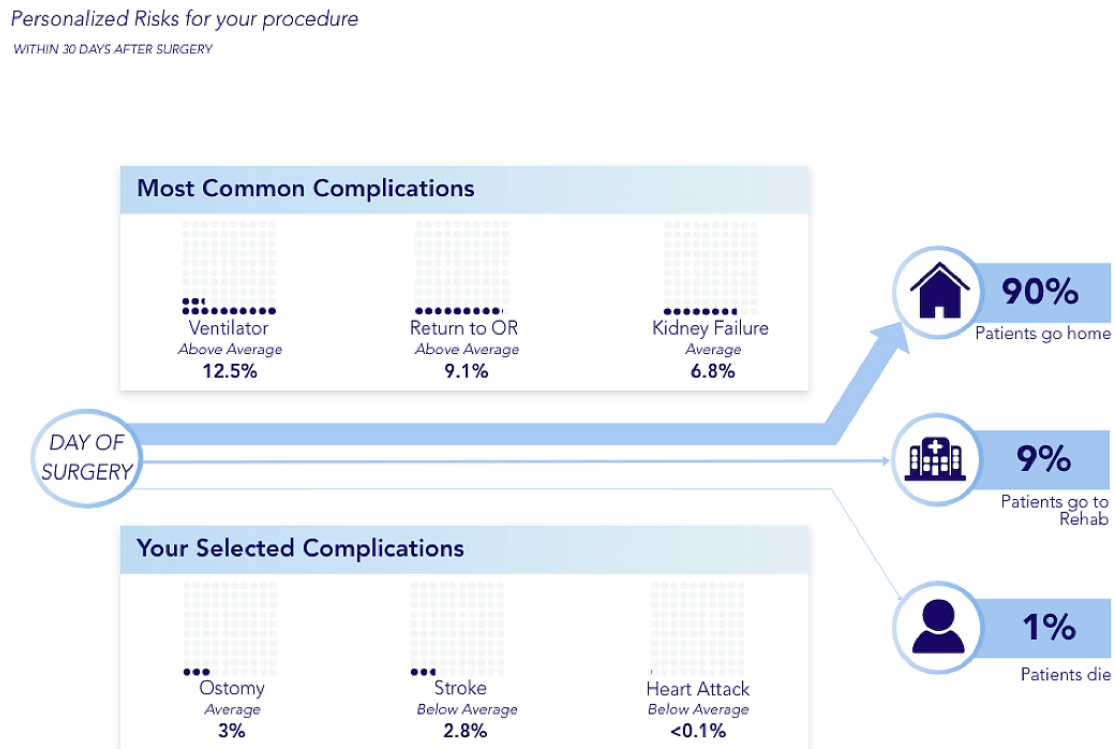
Figure 4. The design of the 3 risk visualizations used in this study. This study investigated 3 visualizations that, through a literature review, were identified as likely to be successful in communicating risks to patients. The top left shows the general layout that all 3 visualizations follow, where the patient’s highest risks are displayed at the top and the complications they have chosen are displayed below to allow for comparison. The likelihood of each discharge destination is separated from the risks and communicated using positional cues and weighted lines. The 3 visualizations tested were bar strength (top right), dot array (bottom left), and logarithmic scale visualization (bottom right). OR: operating room.



We grouped complications into *rare events* (<1%) and *common events* (≥1%). We referred to the Visualizing Health repository [16] to choose visualizations that could be suitable for communicating these events. We chose the *bar strength* visualization that resembles the signal strength on mobile devices and represents a familiar visualization owing to the prevalence of mobile devices (Figure 4, top right). We also chose a waffle chart, called *dot array*, as it is more granular than the bar strength and is recommended by the Visualizing Health repository to accurately communicate risk (Figure 4,

bottom left). To be able to more accurately show risks <1% (compared with the *bar strength* and the *dot array*), we chose a logarithmic scale inspired by the perspective scale proposed by Paling [17], which also allows for direct comparison of risks (Figure 4, bottom right). All of the different visualizations—bar strength, dot array, and logarithmic scale—are shown in the context of the final visualization of the visual consent tool. In Figure 5, we present one instance, the dot array, in a larger image for a better presentation of the design.

Figure 5. An enlarged image of the final stage of the visual consent tool mock-up used for evaluation with one of the possible visualizations—the dot array. OR: operating room.



Evaluation of the Visual Consent Tool

Participants

A convenience sample of 20 patients was interviewed during their postoperative checkup visit to the acute care surgical clinic at an academic hospital. The patients were approached by the interviewer and asked about their willingness to participate in the study. This study only included patients who had undergone a surgical intervention by an acute care surgeon and who agreed to participate with written consent.

Ethics Approval

This study was approved by the Beth Israel Deaconess Medical Center Institutional Review Board (2019P000013).

Interviewer

The interviews were conducted by the first of the 2 joint first authors of this paper (UG). The researcher had no previous relationship with the participants and briefly stated the purpose of the study at the beginning of the interviews. The interviewer did not have any previous biases aside from the assumption that visualization would be a useful tool for the consent process.

Study Procedure

The interview guide was trialed with 2 individuals who were not participants in the study and was refined to fit within 30 minutes and provide answers to our research questions. Figure 6 shows the structure of the semistructured interview (see Multimedia Appendix 1 for the interview guide). The interviews first aimed to understand the patient's informed consent experience in the current practice—without any visualization

aids (Figure 6, part 1 provides answers to research question 1). The second section focused on risk perception and visualization preference (Figure 6, part 2 provides answers to research questions 2 and 3). Finally, the third section assessed perceptions of the value of a visual consent tool and its usefulness during the informed consent process (Figure 6, part 3 provides answers to research question 4). Each of the participants went through the interview only once, and no repeat interviews were conducted.

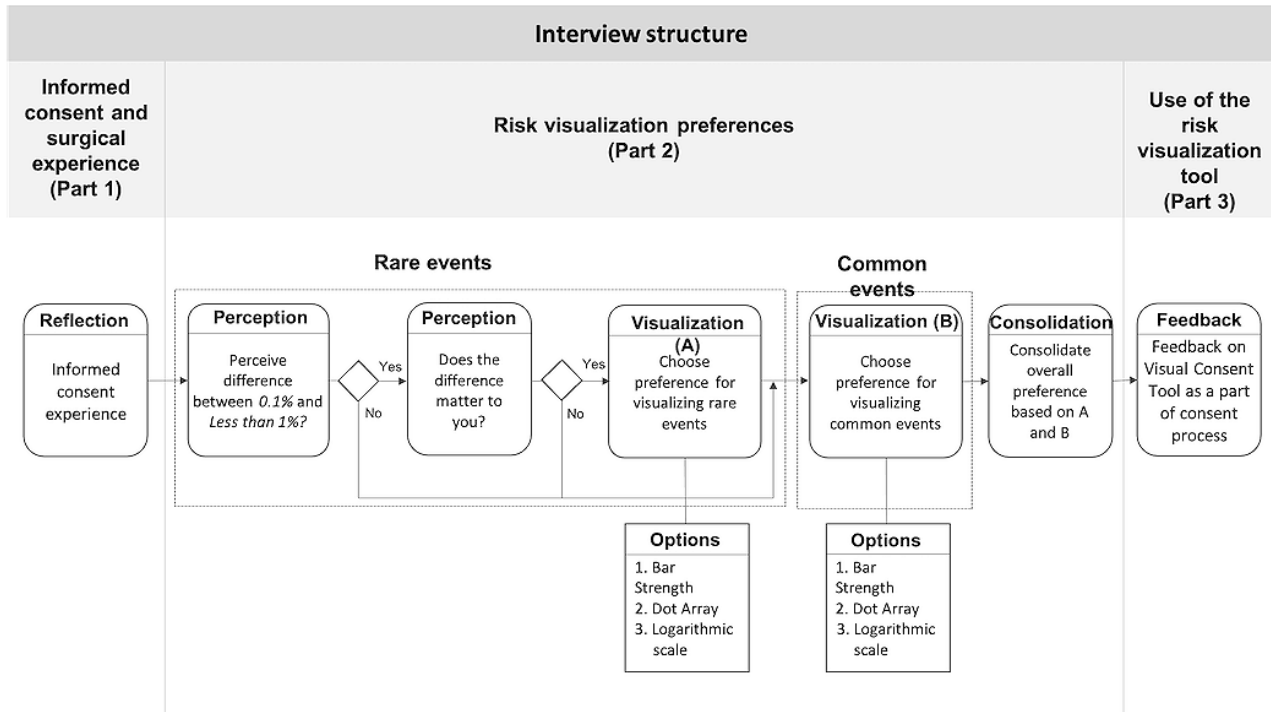
The interviews were conducted in a clinical setting after a postoperative visit. In some cases, the interviews were conducted in the presence of a significant other or family member of the participant. The interviewer gathered demographic data and impressions of the existing informed consent experience by asking the participants to recall their most recent discussion about the risks of informed consent with a surgeon (Figure 6, part 1).

Following this, the interviewer assessed the patient's perception of the risk visualizations (Figure 6, part 2). First, using a broadly familiar example, the interviewer evaluated perceptions of life-threatening rare events by asking the patient if there was a notable difference between 0.1% and the phrase *less than 1%* for a likelihood of being struck by lightning. Patients who perceived a difference were shown 3 visualizations and asked to identify the visualization that conveyed most clearly the 0.1% chance of being struck by lightning. Using a similar approach, the interviewer assessed visualization preferences for more common life-threatening events by using an example of a 12.3% chance of an earthquake. Again, the patient was shown the same 3 visualizations and asked to indicate which one conveyed this

information best. In cases where the visualizations chosen for rare events differed from the visualizations chosen for common events, the patients chose one of the two preferred options for

showing all possible risks: for rare events and more common events.

Figure 6. Semistructured interview structure. Part 1 assessed the current informed consent experience. Part 2 assessed risk perception and risk visualization preferences. Part 3 gathered feedback on the useful elements of the visual consent tool and its applicability.



In the final step (Figure 6, part 3), the final visual consent tool risk and discharge visualization was presented to the patient using the previously selected visualization type. To test the intuitiveness of the design, the patient was asked to explain what they saw and what decisions could be made without receiving any explanation of the final visualization. The interviewer then explained the intended purpose of the visualization and collected additional comments about the visual consent tool.

Finally, the patient was asked to identify situations in which they would find this tool useful, what they found most useful, and what could be improved.

Data Analysis

The interviews were recorded using an iPhone (Apple Inc), and no field notes were taken during the interview. Interviews were transcribed using Dragon Dictate 3.0 (Nuance Corporation) with manual verification by the interviewer. Transcriptions were not sent back to the participants for comments and corrections. We conducted a mixed methods analysis of the data. For the qualitative part, the two joint first authors of this paper (UG and DN) conducted a thematic analysis of the data using Microsoft Word and Excel. Each of them inductively coded selections of the transcribed interviews independently by assigning labels to the meaningful discourse units in the patients' answers. Both researchers reviewed the codes and clarified any disagreements. Codes were collaboratively grouped into categories, and the categories were grouped into themes. These were checked for validity with domain experts and carefully modified to

accommodate feedback. We did not conduct any member checks with participants. We also applied a quantitative approach to the analysis and extracted discourse units that expressed different preferences and reasons for those preferences. We used the discourse units identified for each interview to assess the prevalence of different perceptions of the visualizations presented in our visual consent tool.

Results

Overview

We interviewed 20 patients attending a postoperative visit. The average age of the cohort was 61.7 (SD 14) years with a range of 29 to 87 (median 59) years, with 55% (11/20) female and 45% (9/20) male participants. The education level ranged from *some high school* to *Ph.D.* Most patients (17/20, 85%) had surgery on the intestines, gallbladder, or appendix, and approximately half of the cases (10/20, 50%) involved emergency procedures.

The thematic analysis of the semistructured interviews revealed three main categories: (1) *factors that influence risk perception* (research question 1), (2) *perceptions of the visualizations* (research questions 2 and 3), and (3) *effects of the proposed visual consent tool and use case scenarios* (research question 4).

Factors That Influence Risk Perception

Overview

We found that patients reacted positively to learning that risks could be personalized. Some stated that personalization of risks was the highlight of the tool as it made them feel more considered as patients. We identified several factors that influenced the perception of those personalized risks, which encompass *the surgical procedure, the patient's cognitive state, and the timing of consent.*

The Surgical Procedure

Factors that influenced risk perception of the procedure included the clarity of diagnosis, the complexity of the procedure, and the urgency of the case. We found that most patients (17/20, 85%) had a clear diagnosis and were confident in the surgeon's familiarity with the case. They indicated that they were less concerned about the risks associated with their surgery compared with patients with an unclear diagnosis. The latter group made statements that emphasized the uncertainty of what was about to take place, which increased their anxiety, and made comments such as “[the surgeons] didn't know what they were getting into.”

Similarly, patients who underwent routine procedures were generally less threatened by the risks compared with patients who were supposed to go through a complex surgical intervention that involved multiple subprocedures. For patients who returned to the operating room, all complications were of low importance, and pain or fear of death outweighed all others.

Finally, patients who had an extended time before the surgery were more willing to analyze the risks and discuss them in greater detail with the surgeon. In contrast, patients who had to go through an emergency surgery were less motivated or even incapable of any form of analysis and were mostly focused on their chance of survival.

The Cognitive Capacity of the Patient

We found that the cognitive capacity of the patient, such as the *capacity for unobstructed thinking, medical knowledge, and literacy*, played a key role in risk perception.

Patients who were in pain or feeling drowsy cared less about complications and wanted to proceed with the surgery as soon as possible.

A consistent theme was the delegation of decision-making to a more medically knowledgeable and literate family member such as a spouse or child when such an opportunity existed. Patients with low health literacy were more likely to not understand the diagnosis and felt that identification of complications was of low importance. These patients would completely delegate the decision-making to the medical professional and restrain themselves from engaging in contributing to the process.

The Timing of Consent

The timing of consent varied among patients. For emergency cases, risks were communicated within a few hours of the surgery; for transfer cases or planned operations, risks could have been initially communicated a couple of weeks in advance. For patients with acute conditions, the lists of complications were of low importance as pain and death were described as important factors. Alternatively, patients with subacute conditions (surgeries within 3-24 hours) felt that knowing the risks was important for expectation management.

Interestingly, although knowing about the risks was of varying importance depending on the timing of consent, all respondents were clear that knowledge about the risks would not have influenced their decision to go forward with the surgery.

Factors That Influence Perceptions of the Visualizations

We identified several factors that influenced the perceptions of the visualizations: *preference for rare event communication, preference for risk visualization, and usefulness of comparison with the average.*

Preference for Rare Event Communication

Most patients (16/20, 80%) did not have a preference for how rare events were communicated. Patients who wanted to know the exact percentage <1% (3/20, 15%) preferred the logarithmic representation for rare events.

Preference for Risk Visualization

Table 1 is based on the preferred visualization after the participant was exposed to the visualization options for rare and common events and asked to consolidate their answers in a single visualization. Table 1 shows the preferences that ranged across the visualizations. Of the three available graphics—the bar strength, dot array, and logarithmic scale—there was no consensus on a preferred visualization.

Table 1. Visualization preferences of the patients.^a

Visualization chosen	Patients, n (%)
Bar strength	5 (25)
Dot array	4 (20)
Logarithmic scale	7 (35)
None	3 (15)
Other	1 (5)

^aThe table shows that the patients had differing preferences for optimal visualization for communicating risks in the visual consent tool. Of the 20 participants, 1 (5%) liked a visual aid but preferred a different visualization from the 3 presented, and 3 (15%) did not express interest in a visual aid.

Patients who preferred the bar strength visualization (5/20, 25%) liked the simplicity and clear step increases, which allowed for quick interpretation. Patients who liked this approach felt that it was less complicated than other options. In addition, 5% (1/20) of patients expressed concern over the discretization of the bars, and a few patients felt that it did not show enough information.

The dot array was endorsed by 20% (4/20) of the patients, who preferred its visual organization and felt that it allowed for comparison of ratios of shaded to grayed out dots. These patients found the dot array easy to understand and that it gave “just the right amount” of information. In addition, 5% (1/20) of patients noted that they would waste time counting dots, and another patient (1/20, 5%) felt that it would be hard to compare risks.

Patients who preferred the logarithmic scale (7/20, 35%) felt that it communicated the risks most clearly and allowed for easy comparison and aggregation of risks. Of those 7 patients, 4 (57%) mentioned that they liked the labeling of I in X . However, patients who did not prefer the logarithmic scale found it the most complicated of the 3 options.

Of the 20 patients, 4 (20%) did not respond positively to the visualizations. In addition, 5% (1/20) of patients was dissatisfied with the choices presented, and 15% (3/20) of participants rejected the visual aids and preferred verbal communication or that the decision be left to the physician. These patients were all aged >75 years, which is notably higher than the average age of the other participants.

Usefulness of Comparison With the Average

Most patients (11/20, 55%) expressed indifference to knowing whether their risk was above or below the average. Those who cared about the average stated that it would raise or lower their concern, and some only cared if it was actionable information. Many were not confident in how to include this information in their decision-making process.

Effects of the Visual Consent Tool and Use Case Scenarios

Overall Impression of the Visual Consent Tool

In terms of intuitiveness, most patients (14/20, 70%) found the final visualization intuitive without any context. However, after explaining the context of the tool and the steps leading up to the final visualization, most patients (15/20, 75%) felt able to make decisions with the help of the visualization.

We observed three major effects of the visual consent tool on perceived informed consent: *depth and length of the discussion*, *information retention*, and *risk awareness*.

Depth and Length of the Discussion

Most patients (13/20, 65%) stated that the visual consent tool would have helped them pick up more information or be more confident in their surgical decision. Most patients (12/20, 60%) claimed that the visual consent tool would allow them to have a better understanding of the possible complications and their likelihood.

All the patients (20/20, 100%) agreed that the visual consent tool would help stimulate a deeper discussion with their provider. They claimed it would “help [them] think of new questions [they] hadn’t thought of before.”

However, some patients (4/20, 5%) expressed concern that having this information and new questions may extend the discussion and would take too much of the surgeon’s time.

Information Retention

A couple of patients (2/20, 10%) also felt that the visualizations might help retain information and suggested using it as a reference to consult after consent.

Risk Awareness

Most patients (11/20, 55%) believed that the visual consent tool would make them more aware of potential risks. This made them more confident in their decision to pursue surgery, but most noted that it would not have changed their decision to pursue surgery.

A number of patients (9/20, 45%) noted that it prompted more long-term thinking about what to expect after the surgery and how it would affect not only them but also their families. In addition, the patients expressed concern that this may be too much information for some patients and that it may dissuade them from pursuing a surgery that was in their best interest.

Patients expressed interest in seeing information generally not available in current risk calculators, such as pain level and expected recovery time. Our interviews revealed that the patients were most concerned about their potential health status and whether they would be able to continue normal activity after surgery—including the chances of avoiding an ostomy (Table 2).

Table 2. The major concerns of the patients before surgery.^a

Concern	Patients, n (%)
Ostomy	6 (30)
Health status after the operation	6 (30)
Postoperative plan	5 (25)
Not laparoscopic	4 (20)
Death	4 (20)
Recovery time	3 (15)
General complications	3 (15)
Anesthesia	3 (15)
Pain medication	2 (10)
Life support	1 (5)
Infection	1 (5)
Blood transfusion	1 (5)

^aThis table shows that the patients were most concerned about their potential health status, possible ostomy, and the pain level they could expect after surgery.

Discussion

Principal Findings

Essentially, the main purpose of our visual consent tool is to empower patients in the decision-making process, provide them with a degree of control over what is being discussed and how the information is being presented to them, and give them the sense that their voice is being heard. To achieve this, we aimed for a high level of personalization in the design, allowing the patients to not only select the risks they wanted to discuss in depth but also account for different risk visualizations to choose from. This approach is different from the traditional one that positions the surgeon as the sole driver of the discussion regarding the risks and assumes that there is one risk visualization type that is suitable for all patients [12,13,18].

The aforementioned approach allowed us to obtain broad insights into how to tackle the design of visual consent tools. In contrast to existing studies that focus primarily on barriers to tool adoption by surgeons, patients' perceptions of the material risk communicated by physician-facing risk assessment tools, or the effect of risk visualization on understanding, the study presented here, to our knowledge, is the first to broadly enumerate the requirements and benefits of a personalized visual informed consent tool that incorporates patient-facing risk visualizations and accounts for patient preference for which risks to be visualized and discussed in detail with the surgeon [18-25]. Through our interviews with patients, we elucidated several unique findings that add to the existing literature and inform the present practice of risk communication and the future landscape of personalized risk visualization. First, and perhaps most significantly, the patients did not identify a single preferred risk visualization, and their preferences varied across the 3 visualizations presented. Second, the patients' concerns regarding postoperative adverse outcomes did not align well with the most probable risks offered by the ACS risk calculator we relied on in our visual consent tool or with the other

traditional risk calculators we reviewed. Third, our visual aid was perceived to improve information retention and risk awareness compared with traditional text-only informed consent documents. These findings will be further discussed below in the context of the current literature.

Variable Preference for Risk Visualization

Almost all patients (15/20, 75%) agreed that the visualizations were useful in communicating risk and would be helpful in their decision-making. Notably, there was no single visualization preferred by most patients, and preferences varied across the available visualizations.

It is difficult to assess whether this finding is aligned with or different from previous findings as most of the literature on this topic of study has evaluated preference for a broad range of visualizations over verbal or textual communication of risks. A few studies have shown that tables, icons, and vertical bar charts are generally preferred over other options such as horizontal bar charts [26,27]. However, the visualizations used in those studies did not cover all 3 visualizations compared in our study.

Although the goal of our study was to determine the preference of patients for risk visualizations, we did not measure understanding quantitatively, though several recent studies have demonstrated differences between participant understanding and preference when presented with different graphical formats to communicate health information [28,29].

Presenting the appropriate risk visualization to a given patient is of high importance but also very challenging. Along these lines, researchers have found that allowing patients to choose a preferred visualization that they feel motivated to interact with versus showing them the useful one that will help them in understanding and using the information better is a decision with trade-offs [30]. To corroborate this, the findings from a study indicated that risks presented in the form of random icons and stacked vertical bar graphs may affect the likelihood of choosing surgery or cause patients to view certain risks as more

complex or threatening [31]. Furthermore, a study of 45 adults contemplating the risks and benefits of recombinant tissue plasminogen activator for ischemic stroke concluded that, although patients preferred bar graphs for risk information, accurate recall and confident decision-making decreased when using the bar graph compared with an icon array or stacked graph [29]. In addition, although bar graphs were preferred, patients spent more time studying them compared with the 2 other graphical formats despite these longer decision times correlating with less accurate recall [29]. The question of how to visually present surgical risks to patients is further complicated by our finding that the desire for risk information and involvement in decision-making varies per patient.

In this context, our findings suggest a need for highly tailored patient-facing decision aids with increased flexibility in visualization beyond a one-size-fits-all approach.

Misalignment Between Patients' Concerns and Current Risk Calculators

Notably, we found that patient concerns were discordant with the risks presented by traditional risk calculators. According to the interviewed patients, postoperative pain, changes in overall health status, familial burden, and adverse functional outcomes were additional considerations before undergoing a procedure that were very rarely or almost never discussed with them. In comparison, traditional risk calculators highlight major causes of perioperative morbidity, such as the risk of renal failure or venous thromboembolism. If we are to consider surgical risk calculators as a step toward improved shared decision-making, our interviews suggest that it is important for physicians to leverage these tools to communicate the risk of major changes in quality of life, expected functional outcomes, and consequences of the procedure as they are key tenets of informed consent [22,32]. In some cases, patients may consider these risks to outweigh the clinical consequences when considering whether to pursue surgery and, therefore, these risks should be communicated as well and with equal attention. For example, a study showed that 18% of patients with postopen abdominal aortic aneurysm (AAA) repair would not undergo AAA repair again knowing that the recovery process negatively affects functional activity (such as driving and shopping, among other daily tasks), despite understanding the life-threatening consequences of potential AAA rupture [33].

Although we acknowledge that the aforementioned functional consequences can often be subjective, intangible, and therefore more difficult to capture reliably and at the scale of traditional clinical outcomes, patient demand suggests that these risks should also be prioritized and incorporated into the consent tools. Our tool, for example, addressed the likelihood of the patient returning home compared with the patient not returning home (ie, to a skilled care, acute care, or rehabilitation facility), incorporating an example of a procedure's consequences to the patient's lifestyle. Given patient feedback, and to further mirror the scope of informed consent, future iterations of surgical risk calculators should attempt to explicitly incorporate the probability of additional adverse quality-of-life outcomes and the risks associated with not pursuing a surgical intervention.

Benefits From Using the Visual Consent Tool

The study participants did find benefits from the proposed visual consent tool, which aligns with the understanding that, generally, patient-facing decision aids have numerous benefits for patients [34]. Similar to other studies that assessed patients' desire for risk information, the patients in our study believed that the visual consent tool has the potential to improve information retention and risk awareness [22]. However, they were concerned about how being introduced to a high number of risks might become overwhelming at times, overburden the patient, and maybe dissuade them from going through a surgery that could actually be their best option. Nevertheless, patients using similar surgical risk calculators have reported that preoperative education regarding postoperative risks actually decreases anxiety, with meta-analyses indicating that their use is associated with reduced decisional conflict and increased knowledge [34]. Although some data have shown that patients using decision aids are more likely to choose more conservative or less invasive treatments, other data have shown that the use of surgical risk calculators did not dissuade or discourage patients from pursuing surgical treatment [22,34]. These findings, combined with those of our study, support the idea that delivering risks to patients should be tailored to their needs and preferences. However, determining how many and which risks to show and when requires further research.

Along these lines, and based on the participants' perceptions in our study, we found evidence that our visual consent tool can improve shared decision-making and be beneficial for patients and providers if appropriately customized to the particular context pertinent to the patient. This hypothesis is, of course, subject to further quantitative studies on an updated version of the visual consent tool based on the findings of this study.

Putting the Visual Consent Tool in Broader Real-life Context

Finally, an important consideration are the stakeholders involved in incorporating the visual consent tool into the current clinical workflow. These stakeholders include the patient and their family, surgeons, and hospital administration. This study focuses on the preferences of the patients, but future work should consider input from surgeons and hospital administrators to find a solution that maximizes benefits for all. For example, although the visual consent tool exhibits benefits for the patients, some of them expressed concern that the interactions stimulated by the introduction of the visual consent tool might take too much of the surgeon's time and negatively affect their clinical productivity. Future work should consider how to enable visual consent tool-based communication efficiency that will benefit both patients and surgeons and not significantly favor one over the other. Although the visual consent tool may disrupt current practices, it is also important to consider the greater value of such a tool to the hospital and its administration. Most patients in our study agreed that the visual consent tool can likely raise awareness, stimulate new questions, and allow them to reflect on the discussion with their surgeon after the conversation. As a consequence, the participants believed that these benefits would allow them to take a more active role in their treatment plan. For these reasons, we anticipate that the proposed visual

consent tool will help promote shared decision-making by empowering patients with confidence in their decisions and attenuating the opportunities for miscommunication. Therefore, we can expect that more comprehensively informed patients will be less likely to pursue legal action when they experience a nonbeneficial outcome [5]. We believe that the proposed principles in our visual consent tool and the benefits they could bring show promise not only for patients but also for the health care system as a whole.

Limitations

Our findings should be considered in light of the limitations of this study. The study population was biased toward older patients and only included patients who underwent a specific group of general surgeries. Their relatively positive experiences and historical exposure may have influenced their recall and opinions. Patients who undergo different surgeries, have a different demographic makeup, have worse outcomes, or are in the preoperative period may have different risk perceptions and risk visualization preferences than our study population.

Although we covered only a specific group of general surgeries, we still included a variety of them. This approach may make the results look less focused; however, it was optimal to have a setup that enabled us to learn more comprehensively about the factors that influence risk perception.

Finally, although we understand that the visual consent tool is supposed to be used with preoperative patients, for the purposes of our study, it was actually beneficial to have postoperative patients. The reason for this is that they had a chance to go through the standard consent process and the surgery and assess how different that process should have been. Once presented with the visual consent tool, they were able to evaluate how the visual consent tool might fill in the gaps in the standard consent process based on their experiences.

Conclusions

We found that current risk calculators do not account for a number of concerns patients have, primarily related to their quality of life after the surgery, and suggest that efforts should be made to incorporate these risks into the risk calculators and the consent process. Most importantly, we identified that there is no universal way of visually communicating risks to patients, which counters the current practice of using a single approach. We found that multiple factors affect the perception of risks and that the proposed visual consent tool has the potential to provide useful information to patients and stimulate shared decision-making with their surgeons. We anticipate that these benefits can be achieved if patient characteristics are taken into account to deliver a tailored risk visualization solution. Finally, we postulate that the need for tailored visual communication of complex medical information applies to other domains of health care as well.

Acknowledgments

The authors thank Maura Cass from IDEO for valuable feedback on their evaluation study. They also acknowledge Shannon Ehmsen at Harvard Medical School for help with the graphic design of the visualizations.

Authors' Contributions

UG and DN contributed equally as first authors.

Conflicts of Interest

NG is a cofounder and equity owner of Datavisyn.

Multimedia Appendix 1

Semistructured interview guide.

[[DOCX File, 594 KB - humanfactors_v9i2e29118_app1.docx](#)]

References

1. Surgery 2015-2017 final report. National Quality Forum. URL: https://www.qualityforum.org/Publications/2017/04/Surgery_2015-2017_Final_Report.aspx [accessed 2022-02-25]
2. Cooper Z, Courtwright A, Karlage A, Gawande A, Block S. Pitfalls in communication that lead to nonbeneficial emergency surgery in elderly patients with serious illness: description of the problem and elements of a solution. *Ann Surg* 2014 Dec;260(6):949-957. [doi: [10.1097/SLA.0000000000000721](https://doi.org/10.1097/SLA.0000000000000721)] [Medline: [24866541](https://pubmed.ncbi.nlm.nih.gov/24866541/)]
3. Angelos P. The evolution of informed consent for surgery using the best case/worst case framework. *JAMA Surg* 2017 Jun 01;152(6):538-539. [doi: [10.1001/jamasurg.2016.5654](https://doi.org/10.1001/jamasurg.2016.5654)] [Medline: [28146226](https://pubmed.ncbi.nlm.nih.gov/28146226/)]
4. Bickmore T, Utami D, Zhou S, Sidner C, Quintiliani L, Paasche-Orlow M. Automated explanation of research informed consent by virtual agents. In: *Intelligent Virtual Agents*. Cham: Springer; 2015.
5. Grauberger J, Kerezoudis P, Choudhry AJ, Alvi MA, Nassr A, Currier B, et al. Allegations of failure to obtain informed consent in spinal surgery medical malpractice claims. *JAMA Surg* 2017 Jun 21;152(6):e170544 [FREE Full text] [doi: [10.1001/jamasurg.2017.0544](https://doi.org/10.1001/jamasurg.2017.0544)] [Medline: [28445561](https://pubmed.ncbi.nlm.nih.gov/28445561/)]

6. Scheer A, O'Connor AM, Chan B, Moloo H, Poulin E, Mamazza J, et al. The myth of informed consent in rectal cancer surgery: what do patients retain? *Dis Colon Rectum* 2012 Sep;55(9):970-975. [doi: [10.1097/DCR.0b013e31825f2479](https://doi.org/10.1097/DCR.0b013e31825f2479)] [Medline: [22874604](https://pubmed.ncbi.nlm.nih.gov/22874604/)]
7. Fink AS, Prochazka AV, Henderson WG, Bartenfeld D, Nyirenda C, Webb A, et al. Predictors of comprehension during surgical informed consent. *J Am Coll Surg* 2010 Jun;210(6):919-926. [doi: [10.1016/j.jamcollsurg.2010.02.049](https://doi.org/10.1016/j.jamcollsurg.2010.02.049)] [Medline: [20510800](https://pubmed.ncbi.nlm.nih.gov/20510800/)]
8. Kruser JM, Nabozny MJ, Steffens NM, Brasel KJ, Campbell TC, Gaines ME, et al. "Best Case/Worst Case": qualitative evaluation of a novel communication tool for difficult in-the-moment surgical decisions. *J Am Geriatr Soc* 2015 Sep 17;63(9):1805-1811 [FREE Full text] [doi: [10.1111/jgs.13615](https://doi.org/10.1111/jgs.13615)] [Medline: [26280462](https://pubmed.ncbi.nlm.nih.gov/26280462/)]
9. Taylor LJ, Nabozny MJ, Steffens NM, Tucholka JL, Brasel KJ, Johnson SK, et al. A framework to improve surgeon communication in high-stakes surgical decisions: best case/worst case. *JAMA Surg* 2017 Jun 01;152(6):531-538 [FREE Full text] [doi: [10.1001/jamasurg.2016.5674](https://doi.org/10.1001/jamasurg.2016.5674)] [Medline: [28146230](https://pubmed.ncbi.nlm.nih.gov/28146230/)]
10. Ingraham A, Agarwal S, Jung H, Liepert A, O'Rourke AP, Scarborough J. Patient-centered outcome spectrum: an evidence-based framework to aid in shared decision-making. *Ann Surg* 2018 Dec;268(6):980-984. [doi: [10.1097/SLA.0000000000002466](https://doi.org/10.1097/SLA.0000000000002466)] [Medline: [28922208](https://pubmed.ncbi.nlm.nih.gov/28922208/)]
11. Bertsimas D, Dunn J, Velmahos G, Kaafarani H. Surgical risk is not linear: derivation and validation of a novel, user-friendly, and machine-learning-based predictive optimal trees in emergency surgery risk (POTTER) calculator. *Ann Surg* 2018 Oct;268(4):574-583. [doi: [10.1097/SLA.0000000000002956](https://doi.org/10.1097/SLA.0000000000002956)] [Medline: [30124479](https://pubmed.ncbi.nlm.nih.gov/30124479/)]
12. Meguid R, Bronsert M, Juarez-Colunga E, Hammermeister K, Henderson W. Surgical Risk Preoperative Assessment System (SURPAS): III. Accurate preoperative prediction of 8 adverse outcomes using 8 predictor variables. *Ann Surg* 2016 Jul;264(1):23-31. [doi: [10.1097/SLA.0000000000001678](https://doi.org/10.1097/SLA.0000000000001678)] [Medline: [26928465](https://pubmed.ncbi.nlm.nih.gov/26928465/)]
13. Bilimoria KY, Liu Y, Paruch JL, Zhou L, Kmiecik TE, Ko CY, et al. Development and evaluation of the universal ACS NSQIP surgical risk calculator: a decision aid and informed consent tool for patients and surgeons. *J Am Coll Surg* 2013 Nov;217(5):833-42.e1 [FREE Full text] [doi: [10.1016/j.jamcollsurg.2013.07.385](https://doi.org/10.1016/j.jamcollsurg.2013.07.385)] [Medline: [24055383](https://pubmed.ncbi.nlm.nih.gov/24055383/)]
14. Brennan M, Puri S, Ozragat-Baslanti T, Feng Z, Ruppert M, Hashemighouchani H, et al. Comparing clinical judgment with the MySurgeryRisk algorithm for preoperative risk assessment: a pilot usability study. *Surgery* 2019 May;165(5):1035-1045 [FREE Full text] [doi: [10.1016/j.surg.2019.01.002](https://doi.org/10.1016/j.surg.2019.01.002)] [Medline: [30792011](https://pubmed.ncbi.nlm.nih.gov/30792011/)]
15. Tong A, Sainsbury P, Craig J. Consolidated criteria for reporting qualitative research (COREQ): a 32-item checklist for interviews and focus groups. *Int J Qual Health Care* 2007 Dec;19(6):349-357. [doi: [10.1093/intqhc/mzm042](https://doi.org/10.1093/intqhc/mzm042)] [Medline: [17872937](https://pubmed.ncbi.nlm.nih.gov/17872937/)]
16. Visualizing health homepage. Visualizing Health. URL: <http://www.vizhealth.org/> [accessed 2022-02-25]
17. Paling J. Strategies to help patients understand risks. *BMJ* 2003 Sep 27;327(7417):745-748 [FREE Full text] [doi: [10.1136/bmj.327.7417.745](https://doi.org/10.1136/bmj.327.7417.745)] [Medline: [14512489](https://pubmed.ncbi.nlm.nih.gov/14512489/)]
18. Leeds IL, Rosenblum AJ, Wise PE, Watkins AC, Goldblatt MI, Haut ER, et al. Eye of the beholder: risk calculators and barriers to adoption in surgical trainees. *Surgery* 2018 Nov;164(5):1117-1123 [FREE Full text] [doi: [10.1016/j.surg.2018.07.002](https://doi.org/10.1016/j.surg.2018.07.002)] [Medline: [30149939](https://pubmed.ncbi.nlm.nih.gov/30149939/)]
19. Prasad KG, Nelson BG, Deig CR, Schneider AL, Moore MG. ACS NSQIP risk calculator: an accurate predictor of complications in major head and neck surgery? *Otolaryngol Head Neck Surg* 2016 Nov 22;155(5):740-742. [doi: [10.1177/0194599816655976](https://doi.org/10.1177/0194599816655976)] [Medline: [27329422](https://pubmed.ncbi.nlm.nih.gov/27329422/)]
20. Vosler PS, Orsini M, Enepekides DJ, Higgins KM. Predicting complications of major head and neck oncological surgery: an evaluation of the ACS NSQIP surgical risk calculator. *J Otolaryngol Head Neck Surg* 2018 Mar 22;47(1):21 [FREE Full text] [doi: [10.1186/s40463-018-0269-8](https://doi.org/10.1186/s40463-018-0269-8)] [Medline: [29566750](https://pubmed.ncbi.nlm.nih.gov/29566750/)]
21. Lewis CM, Aloia TA, Shi W, Martin I, Lai SY, Selber JC, et al. Development and feasibility of a specialty-specific National Surgical Quality Improvement Program (NSQIP): the head and neck-reconstructive surgery NSQIP. *JAMA Otolaryngol Head Neck Surg* 2016 Apr 01;142(4):321-327 [FREE Full text] [doi: [10.1001/jamaoto.2015.3608](https://doi.org/10.1001/jamaoto.2015.3608)] [Medline: [26892756](https://pubmed.ncbi.nlm.nih.gov/26892756/)]
22. Raymond BL, Wanderer JP, Hawkins AT, Geiger TM, Ehrenfeld JM, Stokes JW, et al. Use of the American College of Surgeons National Surgical Quality Improvement Program Surgical Risk Calculator during preoperative risk discussion. *Anesthesia Analgesia* 2019;128(4):643-650. [doi: [10.1213/ane.0000000000003718](https://doi.org/10.1213/ane.0000000000003718)]
23. Johnson C, Campwala I, Gupta S. Examining the validity of the ACS-NSQIP Risk Calculator in plastic surgery: lack of input specificity, outcome variability and imprecise risk calculations. *J Investig Med* 2017 Mar 28;65(3):722-725. [doi: [10.1136/jim-2016-000224](https://doi.org/10.1136/jim-2016-000224)] [Medline: [27793973](https://pubmed.ncbi.nlm.nih.gov/27793973/)]
24. Cologne KG, Keller DS, Liwanag L, Devaraj B, Senagore AJ. Use of the American College of Surgeons NSQIP surgical risk calculator for laparoscopic colectomy: how good is it and how can we improve it? *J Am Coll Surg* 2015 Mar;220(3):281-286. [doi: [10.1016/j.jamcollsurg.2014.12.007](https://doi.org/10.1016/j.jamcollsurg.2014.12.007)] [Medline: [25617914](https://pubmed.ncbi.nlm.nih.gov/25617914/)]
25. Etnel JR, de Groot JM, El Jabri M, Mesch A, Nobel NA, Bogers AJ, et al. Do risk visualizations improve the understanding of numerical risks? A randomized, investigator-blinded general population survey. *Int J Med Inform* 2020 Mar;135:104005. [doi: [10.1016/j.ijmedinf.2019.104005](https://doi.org/10.1016/j.ijmedinf.2019.104005)] [Medline: [31901707](https://pubmed.ncbi.nlm.nih.gov/31901707/)]

26. Hildon Z, Allwood D, Black N. Impact of format and content of visual display of data on comprehension, choice and preference: a systematic review. *Int J Qual Health Care* 2012 Feb;24(1):55-64. [doi: [10.1093/intqhc/mzr072](https://doi.org/10.1093/intqhc/mzr072)] [Medline: [22166856](https://pubmed.ncbi.nlm.nih.gov/22166856/)]
27. Kim S, Trinidad B, Mikesell L, Aakhus M. Improving prognosis communication for patients facing complex medical treatment: a user-centered design approach. *Int J Med Inform* 2020 Sep;141:104147. [doi: [10.1016/j.ijmedinf.2020.104147](https://doi.org/10.1016/j.ijmedinf.2020.104147)] [Medline: [32544822](https://pubmed.ncbi.nlm.nih.gov/32544822/)]
28. van Weert JC, Alblas MC, van Dijk L, Jansen J. Preference for and understanding of graphs presenting health risk information. The role of age, health literacy, numeracy and graph literacy. *Patient Educ Couns* 2021 Jan;104(1):109-117 [FREE Full text] [doi: [10.1016/j.pec.2020.06.031](https://doi.org/10.1016/j.pec.2020.06.031)] [Medline: [32727670](https://pubmed.ncbi.nlm.nih.gov/32727670/)]
29. Poirier MW, Decker C, Spertus JA, McDowd JM. What eye-tracking methods can reveal about the role of information format in decision-aid processing: an exploratory study. *Patient Educ Couns* 2019 Nov;102(11):1977-1984. [doi: [10.1016/j.pec.2019.05.021](https://doi.org/10.1016/j.pec.2019.05.021)] [Medline: [31176556](https://pubmed.ncbi.nlm.nih.gov/31176556/)]
30. Pieterse AH, Finset A. Shared decision making-much studied, much still unknown. *Patient Educ Couns* 2019 Nov;102(11):1946-1948. [doi: [10.1016/j.pec.2019.09.006](https://doi.org/10.1016/j.pec.2019.09.006)] [Medline: [31582048](https://pubmed.ncbi.nlm.nih.gov/31582048/)]
31. Timmermans D, Molewijk B, Stiggelbout A, Kievit J. Different formats for communicating surgical risks to patients and the effect on choice of treatment. *Patient Educ Counseling* 2004 Sep;54(3):255-263. [doi: [10.1016/s0738-3991\(03\)00238-6](https://doi.org/10.1016/s0738-3991(03)00238-6)]
32. Newton-Howes PA, Bedford ND, Dobbs BR, Frizelle FA. Informed consent: what do patients want to know? *N Z Med J* 1998 Sep 11;111(1073):340-342. [Medline: [9785548](https://pubmed.ncbi.nlm.nih.gov/9785548/)]
33. Williamson W, Nicoloff AD, Taylor LM, Moneta GL, Landry GJ, Porter JM. Functional outcome after open repair of abdominal aortic aneurysm. *J Vasc Surg* 2001 May;33(5):913-920 [FREE Full text] [doi: [10.1067/mva.2001.115164](https://doi.org/10.1067/mva.2001.115164)] [Medline: [11331828](https://pubmed.ncbi.nlm.nih.gov/11331828/)]
34. Knops A, Legemate D, Goossens A, Bossuyt P, Ubbink D. Decision aids for patients facing a surgical treatment decision: a systematic review and meta-analysis. *Ann Surg* 2013 May;257(5):860-866. [doi: [10.1097/SLA.0b013e3182864fd6](https://doi.org/10.1097/SLA.0b013e3182864fd6)] [Medline: [23470574](https://pubmed.ncbi.nlm.nih.gov/23470574/)]

Abbreviations

AAA: abdominal aortic aneurysm

ACS: American College of Surgeons

COREQ: consolidated criteria for reporting qualitative research

Edited by A Kushniruk; submitted 01.04.21; peer-reviewed by A Sarafi Nejad, S Azer; comments to author 06.06.21; revised version received 24.08.21; accepted 02.10.21; published 29.04.22.

Please cite as:

Gisladottir U, Nakikj D, Jhunjunwala R, Panton J, Brat G, Gehlenborg N

Effective Communication of Personalized Risks and Patient Preferences During Surgical Informed Consent Using Data Visualization: Qualitative Semistructured Interview Study With Patients After Surgery

JMIR Hum Factors 2022;9(2):e29118

URL: <https://humanfactors.jmir.org/2022/2/e29118>

doi: [10.2196/29118](https://doi.org/10.2196/29118)

PMID: [35486432](https://pubmed.ncbi.nlm.nih.gov/35486432/)

©Undina Gisladottir, Drashko Nakikj, Rashi Jhunjunwala, Jasmine Panton, Gabriel Brat, Nils Gehlenborg. Originally published in *JMIR Human Factors* (<https://humanfactors.jmir.org>), 29.04.2022. This is an open-access article distributed under the terms of the Creative Commons Attribution License (<https://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in *JMIR Human Factors*, is properly cited. The complete bibliographic information, a link to the original publication on <https://humanfactors.jmir.org>, as well as this copyright and license information must be included.

Original Paper

A Reference Architecture for Data-Driven and Adaptive Internet-Delivered Psychological Treatment Systems: Software Architecture Development and Validation Study

Suresh Kumar Mukhiya¹, PhD; Yngve Lamo¹, PhD; Fazle Rabbi², PhD

¹Western Norway University of Applied Sciences, Bergen, Norway

²University of Bergen, Bergen, Norway

Corresponding Author:

Suresh Kumar Mukhiya, PhD

Western Norway University of Applied Sciences

Inndalsveien 28

5063 Bergen

Bergen, 5063

Norway

Phone: 47 55 58 58 00

Email: itsmeskm99@gmail.com

Abstract

Background: Internet-delivered psychological treatment (IDPT) systems are software applications that offer psychological treatments via the internet. Such IDPT systems have become one of the most commonly practiced and widely researched forms of psychotherapy. Evidence shows that psychological treatments delivered by IDPT systems can be an effective way of treating mental health morbidities. However, current IDPT systems have high dropout rates and low user adherence. The primary reason is that the current IDPT systems are not flexible, adaptable, and personalized as they follow a fixed tunnel-based treatment architecture. A fixed tunnel-based architecture follows predefined, sequential treatment content for every patient, irrespective of their context, preferences, and needs. Moreover, current IDPT systems have poor interoperability, making it difficult to reuse and share treatment materials. There is a lack of development and documentation standards, conceptual frameworks, and established (clinical) guidelines for such IDPT systems. As a result, several ad hoc forms of IDPT models exist. Consequently, developers and researchers have tended to reinvent new versions of IDPT systems, making them more complex and less interoperable.

Objective: This study aimed to design, develop, and evaluate a reference architecture (RA) for adaptive systems that can facilitate the design and development of adaptive, interoperable, and reusable IDPT systems.

Methods: This study was conducted in collaboration with a large interdisciplinary project entitled INTROMAT (Introducing Mental Health through Adaptive Technology), which brings together information and communications technology researchers, information and communications technology industries, health researchers, patients, clinicians, and patients' next of kin to reach its vision. First, we investigated previous studies and state-of-the-art works based on the project's problem domain and goals. On the basis of the findings from these investigations, we identified 2 primary gaps in current IDPT systems: lack of adaptiveness and limited interoperability. Second, we used model-driven engineering and Domain-Driven Design techniques to design, develop, and validate the RA for building adaptive, interoperable, and reusable IDPT systems to address these gaps. Third, based on the proposed RA, we implemented a prototype as the open-source software. Finally, we evaluated the RA and open-source implementation using empirical (case study) and nonempirical approaches (software architecture analysis method, expert evaluation, and software quality attributes).

Results: This paper outlines an RA that supports flexible user modeling and the adaptive delivery of treatments. To evaluate the proposed RA, we developed an open-source software based on the proposed RA. The open-source framework aims to improve development productivity, facilitate interoperability, increase reusability, and expedite communication with domain experts.

Conclusions: Our results showed that the proposed RA is flexible and capable of adapting interventions based on patients' needs, preferences, and context. Furthermore, developers and researchers can extend the proposed RA to various health care interventions.

(*JMIR Hum Factors* 2022;9(2):e31029) doi:[10.2196/31029](https://doi.org/10.2196/31029)

KEYWORDS

software architecture; adaptive system; IDPT system; health care systems; ICBT; adaptive strategies; personalized therapies; reference architecture

Introduction

Background

Internet-delivered psychological treatment (IDPT) systems are software applications that offer psychological therapies or treatments through the internet. In our study, we focused on treatments based on evidence-based psychological therapy [1]. *IDPT systems* borrow core ideas from learning management systems and other content management systems (CMSs). However, IDPT systems are more inclined toward the health care domain and have a principal perspective of helping patients cope with their psychological problems. IDPT systems have 2 types of content: psychoeducational materials and treatment exercises.

Problems With Current IDPT Systems

We attempt to address 2 problems associated with the current IDPT systems.

First, despite evidence that web-based interventions can be effective means for mental health morbidities, most of the current IDPT systems are tunnel based, inflexible (unable to adapt according to user needs, preferences, and context), and noninteroperable [2-4]. These restrictions cause a high dropout rate; less personalization; and hence, low user adherence [3,4].

Second, IDPT systems targeting different psychological issues (such as depression, anxiety, bipolar disorder, schizophrenia, and others) have many similarities in psychoeducational materials, intervention structures, and assessment techniques. However, because of the lack of standard documentation, established frameworks, and clinical guidelines, several forms of IDPT models exist. As a result, developers and researchers tend to reinvent their versions of IDPT systems, making them more complex and less interoperable [5]. Interoperability in such IDPT systems is essential for exchanging information from one system to another. With the prevalence of ambient intelligence, several Internet of Things devices have been connected to assess, monitor, and guide patients. These devices require communication with each other. In addition, people migrate from one geographic area to another. Consequently, there is a need to share data from one software system to another.

Objective

To address the issues associated with the current IDPT system, we conducted this study with 2 objectives.

The first objective was to create a reference architecture (RA) [6] of an *adaptive IDPT system*, which can personalize the treatments according to patients' needs and support adaptability, interoperability, reusability, scalability, security, and modifiability. *Adaptability* is the ability of a system to accommodate treatments based on patient needs, preferences, and context. *Interoperability* is the ability of a system to exchange information correctly and use the information being

exchanged. We use reusability in two contexts: (1) the ability to use the treatment for other types of mental health care and (2) the ability to use the component of the IDPT system. *Scalability* is the ability of a system to grow with the number of patients, data, or other factors increases. *Security* is the ability of a system to communicate safely when considering malicious attacks. *Modifiability* is the ability to evolve and maintain systems.

On the basis of the proposed RA, the second objective was to create an *open-source framework* [7] that can be used to develop an adaptive IDPT system. The *open-source framework* was created to aim to (1) improve development productivity; (2) facilitate communication with domain experts; and (3) improve the quality of user interfaces (UIs), user interactions, and user experiences.

The Need for Adaptive IDPTs

IDPTs have surfaced and become one of the most commonly practiced and widely researched forms of psychotherapy [8]. The evolution of IDPTs, coupled with the exponential growth of internet access worldwide, has the potential to reshape the landscape of mental health care. Despite the evolution of IDPT, several patients with mental health issues remain untreated [9,10]. Obstacles to receiving treatment for mental health problems include long waiting lists, limited access to therapy and psychiatric medications, perceived stigma of seeking help, and treatment costs [3,10,11]. IDPT systems have been proposed as a solution to bridge this treatment gap. IDPT removes several barriers to traditional face-to-face therapy, which hinders most patients from receiving efficient psychiatric care [12]. The use of IDPT tools can enhance mental health in several manners:

1. IDPT is available and accessible from anywhere through an internet connection [13].
2. The temporal aspects of accessing the treatments can be substantially improved.
3. The scalability of IDPT can drastically enhance the functional capacity of the care [14]; for example, multiple patients can receive treatment at the same time.
4. IDPT makes the treatment cost-effective for individuals who do not have insurance or cannot afford out-of-pocket fees for treatment.
5. IDPT removes the discomfort and stigma-related issues associated with face-to-face approaches [14].

Despite this evidence, most current IDPT systems are not adaptive and have poor interoperability. These restrictions cause a high dropout rate; less personalization; and hence, low user adherence. Hence, there is a crucial need for an intervention system that can help personalize treatments and increase user adherence. Current learning management systems and CMSs are not designed to capture information on mental symptoms, and they do not monitor treatment progress and relevant data. Moreover, they cannot address the need for personalization and interoperability. To address these intrinsic requirements, we propose a new RA and evaluate it by developing an open-source

framework based on the RA. The proposed architecture relies on the user profiling technique for personalization (Figures 1 and 2) and ontological labeling for interoperability (Figure 3).

Figure 1. The figure depicts our proposed model of a data-driven adaptive internet-delivered psychological treatment system. The patients interact with the intervention, and an analytics server captures those interactions. On the basis of the logged data analysis, a process referred to as user profiling maintains an up-to-date user model to provide the adaptive effect.

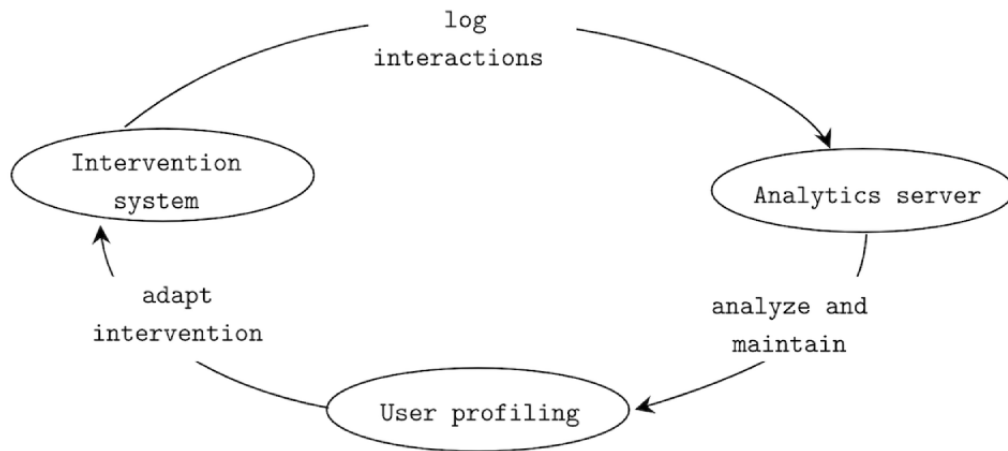


Figure 2. Reference architecture of data-driven adaptive internet-delivered psychological treatment system. AI: artificial intelligence; API: application programming interface; CMS: Content Management System; DL: deep learning; EDA: exploratory data analysis; ML: machine learning; NLP: natural language processing; REST API: RESTful API.

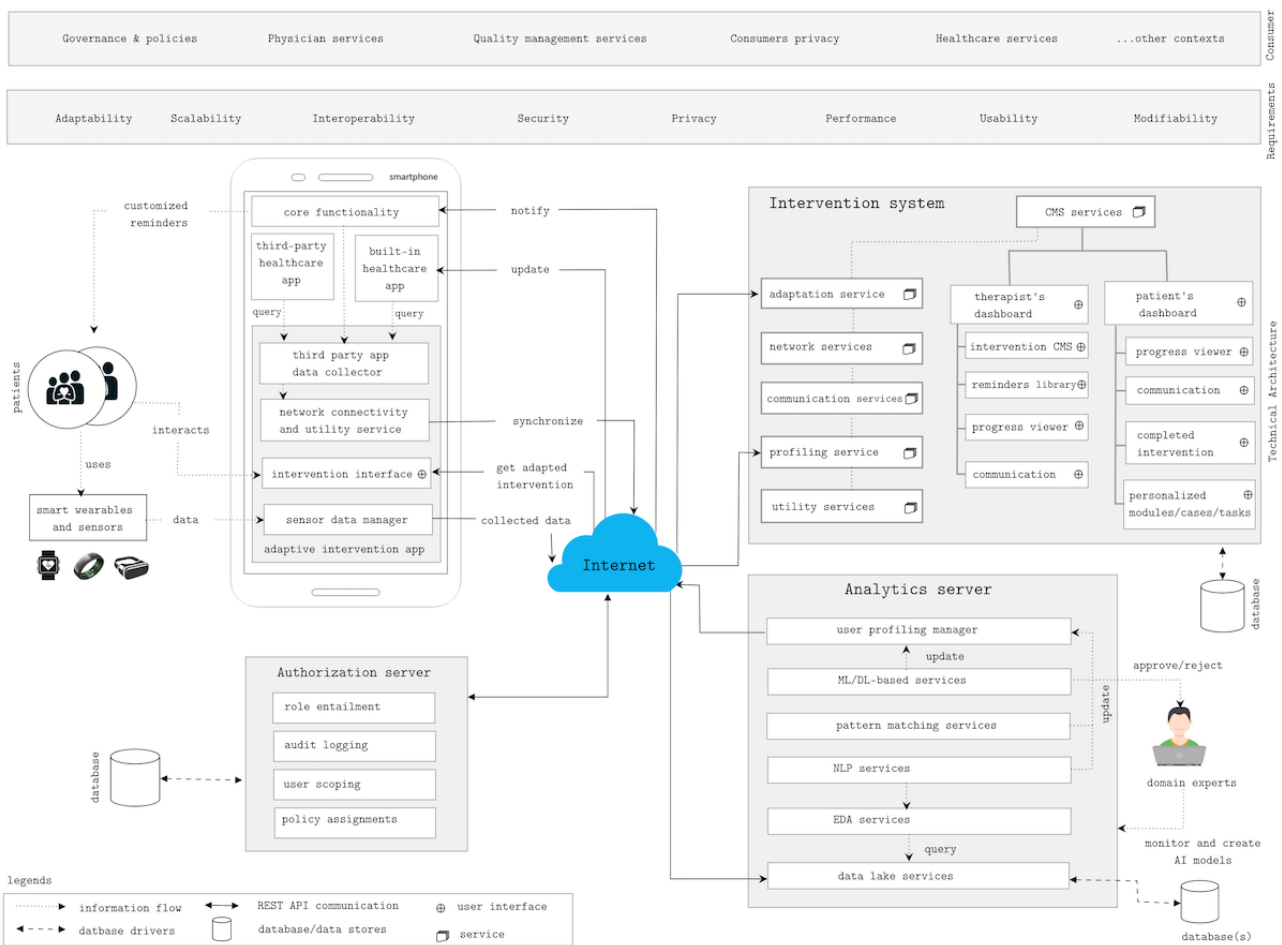
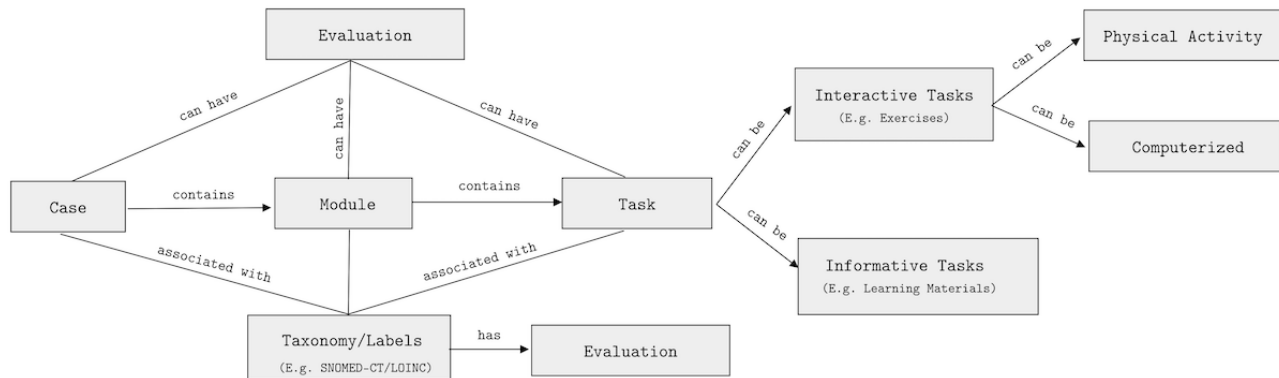


Figure 3. The figure depicts the conceptual model of any intervention that is always associated with a case. A case contains ≥ 1 module. Each module has ≥ 1 task, which can be learning materials or exercises. An exercise can be physical or computerized activities. LOINC: Logical Observation Identifiers Names and Codes; SNOMED-CT: Systematized Nomenclature of Medicine–Clinical Terms.



How We Propose to Increase Adaptivity in IDPT Systems

Figure 1 shows the interaction model of the proposed data-driven adaptive IDPT system. Patients interact with the *intervention system* (see the *Intervention System* section), and an *analytics server* (see the *Analytics Server* section) stores these interactions. In our RA, an analytics server refers to third parties or self-contained services that contain data stores to collect a large amount of data and provide data analytics services such as pattern matching, natural language processing (NLP) service, exploratory data analysis service, machine learning (ML) and deep learning (DL) services among others. On the basis of the analysis of logged data, a process referred to here as *user profiling* (see the *User Profiling* section) maintains an up-to-date user model. A user model is used to provide an adaptive effect. Adaptive systems behave differently for different users. The decision on how the system should behave for any particular user is based on a user model. A user model is a detailed representation of an individual user's information that is associated with an adaptive system. User preferences and needs are dynamic. Hence, it is essential to create, maintain, and update the user model. An adaptive system accumulates data using two distinct approaches to create and maintain an up-to-date user model: (1) implicitly by observing user interactions and (2) explicitly by requesting direct input from the user. This process is referred to as *user profiling*. The essence of the adaptation effect that a system can deliver depends on the nature of the user model's information. Hence, in this study, we aimed to present a framework based on user profiling to provide different adaptation effects.

Contributions

The contributions of this study are 2-fold.

First, we propose an RA for an adaptive IDPT system that provides different adaptation levels based on user profiling. To the best of our knowledge, this novel study is a pioneer in creating, evaluating, and publishing an RA in this domain.

Second, to evaluate the proposed architecture, we created an open-source framework that can be easily extended to several health care interventions. We envision promoting open-source development by creating a proof-of-concept prototype based on the proposed architecture.

Related Work

Grua et al [15] presented an RA for personalized self-adaptive eHealth apps. The proposed RA was envisioned to personalize and self-adapt interventions and increase user engagement with artificial intelligence (AI) applications. The proposed RA uses the Monitor-Analyze-Plan-Execute loop and is primarily targeted at mobile apps. Moreover, the RA follows a client-server architecture and assumes a self-containing, fully flexible AI-enabled back-end system. Such a self-containing back end is neither scalable nor flexible, especially for small-scale health care providers. Health care providers specialize in their domain and dedicate services such as AI, CMS, authentication, and authorization to third parties. For example, in Norway, the health care system relies on a level 4 security system such as BankID [16] for authentication and authorization and AI services from Microsoft Azure or Amazon Web Services. Not all services are coupled into a self-containing system. In such a scenario, *Service-Oriented Architecture* (SOA) is suitable, similar to what we provided in our RA. Unlike their RA, our RA focused on a loosely coupled intervention system that incorporates intervention authorizing services, user profiling services, adaptation services, and others.

WSO2 [17] presents a layered structure that targets scalability and security. On the one hand, the architecture is abstract and domain independent and lacks a specific mechanism to adapt the intervention according to the patient's needs. On the other hand, Wartena et al [18] outlined the RA of a personal telehealth ecosystem referred to as Continua. The proposed RA uses the end-to-end architecture as a design guideline to support interoperability. Continua identifies personal area network devices for communication around a person, local area network devices for communication around a location, wide area network devices for communication around a home and office, and health reporting network devices for communication around enterprise systems such as hospitals, telehealth services, and others. In addition, this study reported how these devices could communicate using associated protocols, promoting interoperability. However, the architecture was abstract, did not address other software quality attributes besides interoperability and security, and did not report how one adapts interventions or personalizes health services.

Rodriguez [19] presented a detailed RA regarding Health care Supportive Homes, a particular type of Ambient Assisted Living (AAL) domain. The proposed RA provides detailed guidelines that can be used to achieve software quality attributes such as interoperability, reusability, security, safety, performance, and reliability. The study reported a detailed and stepwise recommendation for creating a reusable RA. Similar to Continua, Hanke et al [20] presented a *universAAL* reference model for establishing a cross-application platform for AAL. However, both the RAs were specific to the AAL domain. Moreover, they fit the psychological perspective to adapt interventions according to the patient's needs. Mukhiya et al [4] conducted a systematic literature review to identify adaptive elements (content, presentation, feedback message, assessment, activities, reminders, exercises, and reports) of an IDPT system for mental health disorders. The study concluded that most current IDPT systems attempt to adapt feedback messages to patients from therapists. The study reported the lack of an open-source framework for creating adaptive IDPT systems.

Researchers have attempted to theorize user profiling for adaptive web, personalization, and intelligence systems [21-24]. Similar to the studies by Brusilovsky and Millan [21] and Schiaffino and Amandi [24], we considered interest, knowledge, background, goals, individual tasks, and context as essential user profile components. However, in addition to these, we considered several other attributes, including temporal profile, lingual profile, user level, and intervention profile. Furthermore, we modeled these attributes in the proposed framework and illustrated how they could facilitate psychological intervention personalization.

Methods

As a part of the INTROMAT (Introducing Mental Health Through Adaptive Technology) project (see the *Acknowledgment* section), we envisioned developing an adaptive system to offer personalized and customized treatments for patients with mental and neurological disorders. To satiate this goal, we started with the procedures described in the following sections.

Evaluating State-of-the-Art Digital Psychological Treatments Systems Concerning the Current Treatment Requirements

We included usability and universal design principles to evaluate current IDPT systems and publish our findings in this study [25]. Our findings indicated that despite satisfactory treatment results and proven clinical effects, in general, the systems have several issues regarding usability, universal design, and outdated technology.

Collecting Recommendations From Research by Conducting a Systematic Literature Review

We conducted a systematic literature [4] review to (1) inspect and identify the main adaptive elements of an IDPT system, (2) find its information architecture, and (3) determine how adaptation influences the efficacy of IDPT on mental health treatments. The review suggested that adaptive IDPT has the potential to enhance intervention outcomes and increase user

adherence. However, current IDPT systems are tunnel based and do not offer personalized treatment according to user needs. To comprehend how usability is addressed and measured in mobile health interventions for mental health problems, we conducted a systematic literature review [26]. We publish our findings from the perspective of computer science and human-computer interaction in this study [26]. Most studies described their methods as trials, gathered data from a small sample size, and conducted a summative evaluation using a single questionnaire, which indicates that usability evaluation was not the main focus.

Collaborating With Domain Experts and Stakeholders to Comprehend Actual User Needs

Technical domain experts included academicians and industry workers from software engineering, human-computer interaction, AI, and health informatics. Health care domain experts were personal consultants for several mental and neurological disorders. We followed the Domain-Driven Design (DDD) architectural style [27] to model and create the adaptive IDPT framework as these systems involve creating software programs that facilitate the delivery of psychological health care treatments over the internet. Psychological treatments fall under the complex domain, and the development of software systems requires thoughtful collaboration between domain and technical experts. When the domain is complex, it is difficult for designers and developers to build the software. In such cases, developers must steep themselves into the domain to build up their business knowledge. However, most developers do not have much interest in learning about a specific domain in which they are working. In such use cases, the DDD method comes to the rescue.

With the help of domain experts and A technical team, we created the proposed RA. To evaluate the RA, we developed an open-source framework that is presented in this study. We have open sourced the initial prototype under a Massachusetts Institute of Technology License, where everyone is permissible to extend the framework without any consequences. The framework follows the SOA for communication. The server side of the framework follows the Back end for Front end architecture pattern. We followed the Test-Driven Development [28] during framework development. To evaluate the RA's and proposed open-source framework efficacy, we continuously extended the framework for several health care issues, performed randomized controlled trials (RCTs), conducted usability evaluations, and enhanced the system.

Ethical Considerations

This study is a part of the INTROMAT (Introducing Mental Health through Adaptive Technology). As a part of this project, the study was exempted from obtaining ethical approval and the authors have been permitted to publish their findings and research without external approval.

Results

Overview

We use the term *RA* concerning the context and definitions provided by Cloutier et al [6]. As suggested in this study, our

proposed RA encompasses three essential questions: (1) what (ie, the intervention system and its components), (2) why (ie, to adapt and personalize the intervention to enhance user adherence and reduce dropouts), and (3) how (ie, by creating detailed user profiling and using AI and other adaptive strategies to adapt the intervention). One of the reasons for this study is to disseminate the proposed RA, the open-source framework, and ideas for constructing adaptive web applications for health care treatments. To meet the objective of this study, we contextualized a web application at a high level without focusing on specific expertise. From this contextualization, we designed models to comprehend application behavior. Figure 2 shows the RA of the adaptive IDPT system. As the vision is a high level of abstraction, we eliminated constraints related to design, external stakeholders, and others from the model. The RA constitutes 4 major components: *authorization server*, *mobile client*, *intervention system*, and *analytics server*.

Authorization Server

The authorization server [29] is an OpenID Connect-compliant web server [30] with the ability to authenticate patients and grant authorization access tokens. Moreover, the authorization server manages the scopes and permissions of the patients, introspects tokens, entails roles and permissions, audits logs, assigns policies, and requests the intervention system. Our open-source framework included a stand-alone authentication server. However, as the adaptive system follows the SOA architecture, any third-party authentication server can be easily integrated with the framework. For example, in Norway, the use of BankID [16] for authentication or authorization is common.

Mobile Client

The mobile client is the host where the adaptive intervention app (mobile health app) is installed. The mobile client app contains a *third-party app data collector*, *network connectivity and utility service*, *intervention interface*, and *sensor data manager*. The *third-party app data collector* is responsible for communication with third-party apps, health care apps, and built-in health care apps to collect health care data. The *sensor data manager* collects sensors data from Internet of Things devices. The *network connectivity and utility service* are responsible for sending these health care and sensor data to *data lake services* in the analytics server. The mobile client incorporates *intervention interfaces* that allow patients to interact with adapted interventions and communicate with the therapists.

Analytics Server

Conceptually, the analytics server has two parts: (1) the structural part of building a user profile and (2) the analytics method of feeding information to the profile. The analysis servers (Figure 2) incorporate analytical software as a service application programming interface (API). These services take the data as input, detect patterns, and provide a detailed analysis. Although there were several possibilities for the types of algorithms used for data analytics, in the adaptive IDPT context, we aimed for the following core functionalities:

1. *Data lake services* accumulate both sensors and intervention data.

2. *Exploratory data analysis services* help with data cleaning, preparation, exploration, and visualization.
3. *NLP services* help in building, evaluating, and detecting patterns in the textual data set. For example, when patients interact with an intervention, they write some texts as part of computerized exercises. These texts exhibit keywords that express the patient's current state or emotions. It is possible to send these texts directly to available NLP APIs such as Google NLP, obtain the sentiment and tone of the texts, and detect the presence of depression-related keywords [31]. In this study [31], we demonstrate how we can exploit the NLP technique to extract depression symptoms from patient-authored texts.
4. *Pattern matching services* can reveal several associations, correlations, and hidden patterns in the sensor and intervention data. For example, Sharma et al [32] presented a large-scale analysis of the engagement patterns of 35 million posts on 2 popular web-based mental health platforms: TalkLife and Reddit. This study demonstrates that the proposed framework of the engagement patterns enables informative evaluations and analysis of web-based support platforms.
5. *ML and DL-based services* constitute ML and DL algorithms. Data from sensors and interventions can be used to predict early dropout rates and personalize interventions. Once a model has been developed, trained, and evaluated, domain experts can evaluate it for approval. For example, Bremer et al [33] outlined the use of ML techniques to predict dropout in insomnia interventions. Similarly, Nemesure et al [34] proposed an ML approach to predict the presence of generalized anxiety disorder and major depressive disorder.
6. The *user profiling manager* is envisioned to use the analyses and predictions made by NLP services, pattern matching services, and ML and DL services to build a comprehensive profile for the patient.

Intervention System

Overview

The intervention system (Figure 2) is ≥1 web application communicating via *web services* such as RESTful API or GraphQL API. This comprises several services. *CMS services* facilitate a therapist's dashboard (intervention creator, reminder library, progress viewer, communication channels) UI and a patient's dashboard UI (progress viewer, communication channel, history, next or upcoming intervention modules, or tasks). *Adaptation services* provide a rule-based engine for building adaptation rules based on user profiles. *User profiling services* maintain the user profiles. Similarly, *communication services* create a communication channel between patients and therapists. *Network and utility services* handle internet connectivity logic and other utility-oriented tasks. One may separate these components and communicate using a microservice architecture [35]; however, the intervention system components are monolithic for this open-source framework. Figure 3 depicts a conceptual model of the intervention. An intervention is a psychological treatment or therapy delivered through IDPT systems. An IDPT system refers to software that facilitates the creation and delivery of and interaction with

psychological therapy through the internet. These include web applications, mobile apps, augmented reality, and virtual reality applications.

Components

Interventions generally comprise cases, modules, tasks, and taxonomies (labels). In this section, we explain these components, their underlying assumptions (Textbox 1), and their constraints.

Textbox 1. Different assumptions that were considered when designing the open-source framework.

Components and assumptions

Case

- A case contains at least one module.
- A case can have ≥ 1 evaluation criterion.
- A case can have user inclusion and exclusion criteria.

Modules

- A module contains at least one task.
- A module can belong to ≥ 1 case.
- A module can depend on other modules.
- A module can have ≥ 1 evaluation criterion.

Tasks

- A task can have subtasks.
- Each task can have ≥ 1 evaluation criterion.
- The evaluation criteria of a task are the overall evaluation of the subtasks.
- Tasks can have dependency but cannot have their own dependency.

Cases

Typically, IDPTs target ≥ 1 case such as depression, social anxiety, bipolar disorder, attention-deficit/hyperactivity disorder,

Figure 4. Example data structure of a case.

```

CASE_EXAMPLE
├── title: Depression
├── taxonomy: Depression disorder
│   ├── SNOMEDCT: 160329001
│   └── LOINC: 45666-5
├── evaluation: psychometric tests
│   ├── test1: PHQ9
│   └── test2: MADRAS-S

```

Modules

Each case contains ≥ 1 module that focuses on any particular dimension of the case. For example, in the case of depression, there can be modules for understanding and monitoring emotions, behavioral activation, identifying automatic thoughts, and others. A specific module can be part of ≥ 1 case. The modules can have dependencies that specify their ordering.

Tasks

In turn, each module can include ≥ 1 task. A task can be learning materials (informative task) or an exercise (interactive task). *Informative tasks* provide learning materials on mental health issues (cases), symptoms, use cases, and several ways of managing them. The main objective of such informative tasks is to provide psychoeducation so that patients and their families can learn about symptoms, causes, and treatment concepts;

or other health issues. An example of a *case* is shown in Figure 4.

patients can comprehend the self-help program and steps required to manage their illness; and patients can correlate their situations with others who have similar issues, which helps to ventilate their frustrations.

These informative tasks are in the form of reading (text), listening (audio), graphics, presentations, and watching (video). In contrast to informative tasks, *interactive tasks* involve user interaction, often in the form of exercises and psychometric tests. These exercises can be *physical activities* or *computerized tasks*. Examples of physical activities include physical workouts and mindfulness exercises such as breathing exercises, walking certain distances, stretching, or physically performing other activities. Examples of computerized exercises involve filling in blanks, answering (questions and answers), multiple-choice questions, and feedback. The feedback forms comprise the use of free text, rating systems, or multiple-choice questions. The

minimal data structure of a task and its types are illustrated in [Figure 5](#).

Figure 5. Data structure of a task.

```

TASK_DATA_STRUCTURE
├─ ID: Integer
├─ prerequisite: [{ID, score}] .. ARRAY OF IDS
├─ description: String
├─ tags: [String]
├─ points: Integer
├─ completionRequired: Boolean
├─ complexityLevel: Integer
├─ nextTask: ID
├─ type: String == TEXT
│   ├── isRead: Boolean
│   ├── totalReadTime: Integer
│   └─ content: String
├─ type: String == AUDIO
│   ├── isPlayed: Boolean
│   ├── audioLength: Integer
│   ├── playedLength: Integer
│   └─ file: File
├─ type: String == VIDEO
│   ├── isPlayed: Boolean
│   ├── videoLength: Integer
│   ├── watchedLength: Integer
│   └─ file: File
├─ type: String == EXERCISE
│   ├── type: [QUIZ, FEEDBACK, RATING]
│   ├── passingScore: Integer
│   └─ owner: [ID].....BELONGS TO MODULE/TASK
├─ taxonomy: [{key: value}]
└─ evaluation: [{key: value}]

```

Taxonomy or Labels

Each case, module, and task is associated with a label or taxonomy. As cases, modules, and tasks form the hierarchical structure, these taxonomies provide ontological structures for adaptation.

Constraints

A task or module may have ≥ 1 constraint. These constraints determine the states (see the *States of Intervention Components* section) of the task and module. As illustrated in [Figure 5](#), a task or module can have the following constraints:

1. Prerequisite, which is a list of tasks required to be completed before the task is active
2. Next task, which is a task that can have a restrictive follow by selecting the next available task

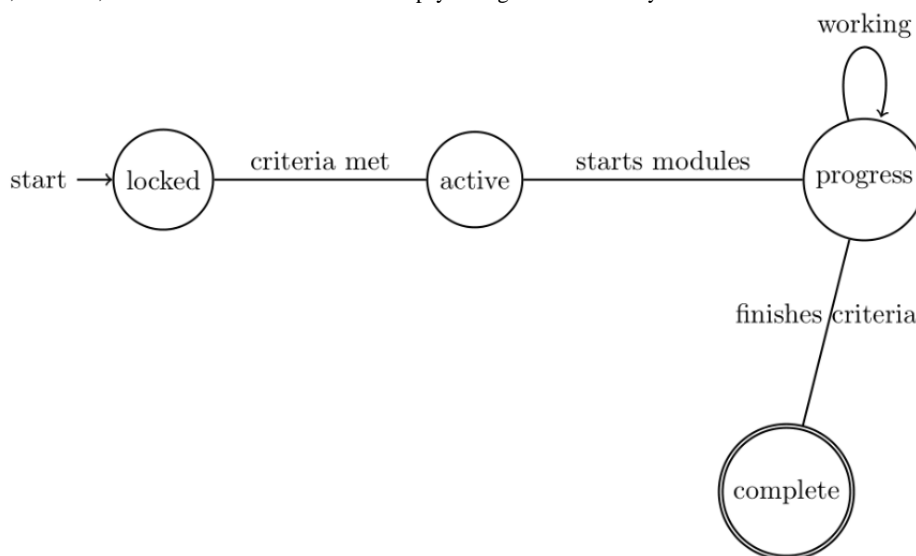
3. Completion required, which is a Boolean value that represents that a user must complete the task if truthy
4. A passing score on an exercise of a quiz type determines whether an exercise is complete

States of Intervention Components

All the cases, modules, and tasks can have four different states, as shown in [Figure 6](#):

1. *Locked*: An entity is locked if its evaluation criteria are not fulfilled or if a dependent entity is not completed.
2. *Active*: An entity is active as soon as the evaluation criteria are matched or its dependent entity is completed.
3. *Progress*: An entity is in progress if it is active, but all the evaluation criteria have not been completed.
4. *Complete*: An active entity is marked as complete if all evaluation criteria are completed.

Figure 6. States of cases, modules, and tasks of an internet-delivered psychological treatment system.



User Profiling

Overview

A profile is a description of an actor containing the necessary facts about the individual. In an adaptive IDPT context, a user profile (or user model) holds essential facts about an individual patient. The process of inferring unobserved data about users from observable data about them (ie, their actions or interactions) is referred to as user profiling [36]. The primary motivation for building user profiling is that users differ in their preferences, interests, backgrounds, goals, cognitive skills, and other attributes. Discovering these differences is essential for presenting users with personalized or adapted services. In an adaptive IDPT system context, user profiling aims to provide an adaptation effect; that is, to behave differently for different users [21]. As mentioned previously, we envision applying user profiling as a fundamental basis for adaptation. Hence, we discuss user profiling in this section and discuss how we can

use such profiling to adapt interventions in the *Scenario-Based Evaluation for Adaptation Support* section. It is essential to note that user profiling can be based on a distributed architecture. Hence, the data-driven adaptive system presented in Figure 1 follows the SOA. An adaptive system tends to find the most relevant information to the user’s interests and presents the information in the right form so that the user may perceive its relevance. A user profile typically powers the discovery of such relevancy and its ranking. A user profile can contain several components (see the *Components of a User Profile* section), such as user interest, knowledge, background, goal, individual traits, and user context. An adaptive system can create and maintain the user profile explicitly and implicitly. We discuss these data acquisition methods in the *Methods of Collecting Information for the User Profile* section. Table 1 summarizes the different aspects of a user profile, including the components of a user profile, the form of representation in software, and the types of data stored in each of the component processes of obtaining data.

Table 1. Different components of use profiling techniques.

Attributes	Common representations	Types of data	Profiling approach
Interest	Weighted vector of keywords; topic hierarchies	News topics, webpage topics, document topics, work-related topics, and hobbies	Implicit or explicit
Knowledge	Scalar modeling; overlay modeling	Application domain	Implicit
Background	Stereotype modeling	Profession, job responsibilities, experience of work, and specific view on the domain	Explicit
Goals or tasks	Goal catalog approach	Goal of the work, information need, and learning goal	Explicit
Individual traits	Mixed approaches	Cognitive styles, personality traits, learning styles, and demographic	Implicit or explicit
Context	Set of name-value pairs	Platform, location, physical environment, social context, and affective state	Implicit

Components of a User Profile

Overview

The content of the user profiles varies according to the system’s domain and the software architect who designed the system.

There are no specific standards that specify which components should be in a profile. Similar to the studies by Brusilovsky and Millan [21] and Schiaffino and Amandi [24], we categorized the user profile content into the components described in the following sections for our framework.

Interests

User interests affect their adherence to software systems. Hence, capturing user interests and attempting to personalize content based on their interests can be an effective means of boosting user adherence. A software system can represent user interest in two ways:

1. The weighted vector of keywords: For example, Lieberman et al [37] used term frequency and inverse document frequency to model user interests. In the term frequency and inverse document frequency technique, each word's weight is computed by comparing the word frequency in a document against the word frequency in all documents in the corpus.
2. Topic hierarchies: A graph can express topic hierarchies where a node is a set of topic words representing a user's specific interest. These types of representations are essential when modeling user interests and associated subtopics.

Knowledge

The user's knowledge represents their understanding of the subject or domain. The user's knowledge is a dynamic feature that increases or decreases over time. Therefore, a well-adaptive system should recognize a user's current state of knowledge and tailor the user model accordingly. A software system can represent a user's knowledge in two ways (Table 1):

1. Scalar modeling: Scalar modeling systems use quantitative scales (for example, 0 to 10) or qualitative scales (eg, excellent, very good, good, bad, poor, and none). However, formulating scalar values for user knowledge is challenging. Hence, scalar modeling has low precision.
2. Overlay modeling: In overlay modeling, the domain contains ≥ 1 subfragment. For each fragment, an overlay model stores the estimation of the user knowledge. The estimation can be binary (knows or does not know), qualitative (excellent, very good, good, bad, poor, or none), or quantitative.

Background

The user's background constitutes information about their profession, job responsibilities, work experience, and a specific view of the domain. The most common representation format for a user's background is stereotype modeling, as detailed background information is not essential. In stereotype modeling, a domain expert distinguishes the most common categories of users according to their background information and adapts the content presented to the user category. The system can also differentiate users by profession (student, medical person, teachers, and others), which implies both knowledge and responsibilities. Several adaptive systems use background information to adapt the content based on the background information of the user.

Goals

The user's goal represents the purpose that the user desires to achieve from the system. The purpose can be information needs, learning goals, or the working of the applications (Table 1). These user goals are dynamic and change over time. Hence, it is essential to tailor the intervention according to the current user's goal. The most common way of representing a user's

goal is to use the goal catalog approach, in which the system presents a predefined set of possible user goals. An adaptive system can recommend certain pages to the user based on a predefined set of goals [21] or adapt the content selected page [38].

Individual Traits

The user's traits include cognitive styles, personality traits, learning styles, or demographic data. Several researchers have agreed on the importance of individual traits and their use in adaptation. Individual traits are stable features of a user, do not change at all or change over a long time, and can be extracted through specially designed psychological tests. Although *cognitive styles*, *personality traits*, and *demographic data* have been discussed in the literature, *learning styles* have been argued [39]. Various methods have been used to extract a user's personality traits and cognitive styles and use them for adaptation.

Context

The prevalence of ubiquitous computing has attracted several researchers of the user's context, such as location, social context, physical environment, and affective state, to tailor software systems. Most of the work on user context has focused on user platforms. For example, most of the studies attempted to adapt to make the system responsive [40] or tailor the content based on hardware, available software, and bandwidth. Affective contexts include physiological and mental contexts. The social context comprises the current user's social aspects, such as information about friends, neutrals, enemies, neighbors, coworkers, and relatives. The most common way of storing the user context is in the form of a *key-value* pair.

Methods of Collecting Information for the User Profile

As mentioned previously, there are 2 ways of extracting the information required to build a user profile: *explicitly* or *implicitly*.

Explicit Information Extraction

A software system can extract profile components such as backgrounds, goals, and interests explicitly; that is, by asking users through UIs such as forms or feedback. Generally, users are not willing to fill in long forms to provide information about them; hence, they are optional. The information accumulated in this way includes demographic data such as age, job, and hobbies.

Implicit Information Extraction

Explicit method of user information has several challenges, including (1) users do not like to fill up long forms, (2) users do not always tell the truth when made obligated to fill forms, and (3) users who wish to fill up the form willingly may not know how to express their interests in words. Observing user interactions (time spent on the content page, bookmarked pages, amount of scroll, content viewed, video watched, and others) with a software system and logging these actions, we can obtain information about users through ML or data-mining techniques. A vital advantage of the implicit method is that we can log and analyze users' changing interests, preferences, habits, and goals

over time. These logs can help adapt the content or presentation according to the correct context of the user.

Discussion

Principal Findings

Architectural evaluation ensures that the architectural design decisions produced are the correct ones [41]. One of the RA evaluations aimed to analyze and verify that it addressed the problems identified in the current IDPT systems. We chose both empirical (case study) and nonempirical (expert evaluation and scenario-based method) evaluation techniques to analyze and verify the proposed RA and open-source framework. Two relevant options for scenario-based evaluation are the software architecture analysis method (SAAM) [42] and the architecture trade-off analysis method [43]. We chose to apply the SAAM method as we proposed RA and qualitatively evaluated it. Moreover, SAAM is suitable for assessing whether a given RA satisfies a specific system's desired properties, whereas the architecture trade-off analysis method is more suited to determine the trade-off between architectural alternatives [43], as performed in the *Related Work* section.

Scenario-Based Evaluation for Adaptation Support

Overview

The first SAAM method was to develop scenarios. This section presents how the proposed RA can perform different types of adaptations based on user profiling. These scenarios, along with the open-source framework's initial prototype, served as scenarios for our evaluation process. The next step in SAAM is to describe the candidate RA outlined in the *Discussion* section. After that, it is necessary to identify the system quality attributes with the help of the developed scenario. The Software Quality Metrics section describes the identified software quality attributes extracted from our scenarios.

Content Adaptation

Content adaptation may involve two subcategories: (1) content materials adaptation, which involves deciding what content is the most relevant to the current user, and (2) content presentation adaptation, which involves determining how to present the selected content effectively to the current user.

Content Materials Adaptation

The main task was to identify the most relevant content for a given user in their context and how to organize that content.

Figure 7. An example scenario of human emotion taxonomy.



User profile components such as *interests, preferences, background, knowledge, and goals* can help to select the most relevant content for a given user. The literature mentions 2 different approaches for adapting content: *page variants* and *fragment variants* [44]. Our RA supported both approaches for content adaptation. Moreover, with user profiling in place, we can perform *adaptation based on the metadata*. The details of the approaches are as follows:

1. In the *page variants approach*, different versions of each page are created using the CMS service (Figure 2). The chosen adaptive strategy [3] selects and presents the most suitable content to the user based on its current context and profile.
2. A page is divided into ≥ 1 fragment in the *fragment variants approach*, where each fragment corresponds to a self-contained element such as text, audio, video, paragraph, picture, or presentations. In an IDPT system, these fragments are authored by domain experts. These fragments are selected and presented to the user based on an appropriate adaptive strategy. Currently, most IDPT systems use rule-based adaptation mechanisms to predefine these fragments. However, once these fragments are predefined and labeled correctly, many pages can be automatically generated to cover a correspondingly large number of interaction contexts.
3. As illustrated in Figure 5, in metadata or taxonomy-based adaptation, a task can have several tags. These tags act as a list of controlled vocabulary sets that define several dimensions of a text. For example, a text can provide psychoeducation about different human emotions (*sad, happy, angry, disgust, sadness, joy, love, and surprise*). These controlled vocabularies can be abstracted to form a taxonomy and ontology related to any particular illness (Figure 7). For example, *joy* and *love* indicate positive emotions. Similarly, *anger, disgust, sadness, and fear* indicate negative emotions. While a user reads a text about *disgust*, these taxonomies can help recommend other tasks (audio, video, images, or activities) that exhibit similar labels as *anger, sadness, or fear* as they indicate negative emotions at a higher level. Moreover, *emotions* are associated with other mental health issues. Learning materials related to emotions for *depression cases* can also be used in *social anxiety cases*.

Content Presentation Adaptation

Let us assume that a task can be represented by the following modalities: audio, video, slides, and text. Here, we assume that each format preserves the semantic meaning of the original format. An IDPT system can personalize a content format based on user interests (see the *Interests* section) and goals (see the *Goals* section). An IDPT system can obtain user interests based on (1) a process mining technique, (2) explicitly asking users, (3) user interaction data, and (4) other data-mining techniques. Process mining can reveal the format of the content that a particular user interacts with the most. If a user spends more time watching videos, the system can present the next video format task. In addition to process mining, user interaction data can reveal the preferred content format for any particular user. These preferences and interests are stored in the user profile and are used for content format adaptation.

Reminders or Alert or Other Notifications Adaptation

Figure 8 illustrates an example of different tasks inside a module in a typical IDPT. As shown in the figure, *Task 2* and *Task 3* have *Task 1* as a dependency. This dependency means that a patient must finish *Task 1* before *Task 2* and *Task 3* are active for them. In addition, to complete *Task 1*, the patient must fulfill both evaluation criteria *E1.1* and *E1.2*. Once *Task 1* is completed, it is marked *complete*, and *Task 2* and *Task 3* are *active*. In addition, the IDPT system schedules an automatic alert or notification for the patient with a personalized message indicating the completion of the task and availability of the next tasks. As shown in Table 2, the process of customized alerts or notifications can be adapted based on the task status. The alert or notification adaptation scenario presented here is an example and can be extended based on other criteria.

Figure 8. An example of different tasks inside a module in an internet-delivered psychological treatment system. Here, the dotted arrow denotes dependencies. For example, Task 2 is dependent on Task 1. Such dependency indicates that a patient cannot start Task 2 before Task 1. Each task has ≥1 evaluation criterion denoted as E.x.

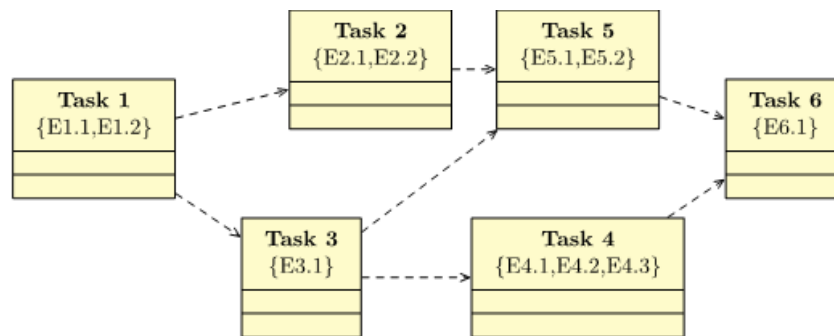


Table 2. An example illustrating alert or notification adaptation.

Task and evaluation	Completion	Notifications
Task 1		
E1.1	✓ ^a	<ul style="list-style-type: none"> Task 1 completion alert Task 2 and Task 3 are active
E1.2	✓	<ul style="list-style-type: none"> Task 1 completion alert Task 2 and Task 3 are active
Task 2		
E2.1	✓	<ul style="list-style-type: none"> Task 2 completion alert
E2.2	✓	<ul style="list-style-type: none"> Task 2 completion alert
Task 3		
E3.1	✓	<ul style="list-style-type: none"> Task 3 completion alert Task 4, and task 5 are active
Task 6		
E6.1	✓	<ul style="list-style-type: none"> Task 6 completion alert Module 1 completion SMS text message

^a✓ indicates that a user has completed the task.

User-Level Adaptation

As depicted in Figure 5, each task has an evaluation (*points*) associated with it. Once a user completes the task, the user

obtains these points. The sum of the points obtained from each task indicates the overall score for any user. An adaptive system can have a simple adaptive rule to activate or deactivate tasks

based on their overall score. For example, if we know the total score of $T_{overall}$ for a user, we can activate or block the availability of the next task for that user. We can use a simple rule engine, such as the following, to activate or deactivate tasks:



In the above rules, T_i (T1, T2, and T3) is a list of tasks. In the above example, we assumed that the threshold score for each task could be decided empirically or determined by the therapists who designed the intervention. According to this example, if $T_{overall}$ is between 0 and 40, we would recommend Task 1 to the patient. Similarly, if $T_{overall}$ is between 41 and 80, we would recommend Task 2 to the patient.

Software Quality Metrics

Overview

As previously mentioned, we envisioned addressing the challenge of high dropout and low user adherence in the current IDPT system. Therefore, a primary software quality metric based on International Organization for Standardization/International Electrotechnical Commission (ISO/IEC) 25000 [45] is adaptability. Moreover, based on the current IDPT system analysis, recommendations from the literature review, and discussion from domain experts, our secondary software quality attribute requirements include scalability, interoperability, security, reusability, and modifiability. We have adopted the notational convention keywords *MUST*, *MUST NOT*, *REQUIRED*, *SHALL*, *SHALL NOT*, *SHOULD*, *SHOULDN'T*, *RECOMMENDED*, *MAY*, and *OPTIONAL* in this section to describe these software quality attributes and compliance in the proposed RA. These keywords are to be interpreted as described in Request for Comments (RFC) 2119 [46].

Adaptability

We aimed to adapt interventions according to user needs and requirements to enhance user engagement and increase adherence. To adapt the intervention, we created a detailed user profile. On the basis of these profiles, we adapted the intervention. The discussion section provides several scenarios explaining how the proposed RA fulfils this need.

Scalability

The entire data-driven adaptive system is based on the SOA. The SOA enforces scalability by organizing services into several components that communicate over a network. Each component of the architecture can be updated and evolved in terms of hardware and software, independent of other components. For example, the IDPT intervention system server's data storage capacity can be increased or decreased without affecting the analytics server.

Interoperability

Our framework supports taxonomic labeling. These are the basics of ontology. On the basis of these taxonomies, we can define several ontology codes such as *Systematized Nomenclature of Medicine—Clinical Terms*, *Logical Observation*

Identifiers Names and Codes (LOINC), and others. The support for such taxonomies will allow us to gain interoperability. To enforce interoperability, we used *Health Level Seven International Fast Healthcare Interoperability Resources (HL7 FHIR)* as the underlying communication standard.

Reusability

The proposed RA uses an SOA that supports reusability to a great extent. For example, we can use the authorization server to handle authentication and authorization for several services. We can reuse the interventions for other health care treatments. Similarly, the analytics server was loosely coupled with the RA and can be reused for several different data analysis purposes.

Security

The authorization mechanism *must* be Transport Layer Security secured [47] and should be improved using the contemporary practices mentioned by the Internet Engineering Task Force [47]. For the prototype, the authorization is incorporated inside the IDPT system but is subject to change as a separate SOA component, similar to the authorization server used [13]. In any case, the authorization *should* issue short-lived tokens and have a mechanism open to administrators and end users to eliminate tokens in the case of a security conflict.

Modifiability

Modifiability incorporates *evolvability*, *customizability*, *configurability*, and *extensibility* [13]. The SOA-based architecture facilitates modifiability by allowing the manageable growth of systems [48]. These systems and components are independent of vendors, products, and technologies. This independence makes it easy to manage and modify individual components. For example, the analytics server in the architecture (described in the *Discussion* section) *may* update the ML libraries or create an additional service that consumes data and performs business intelligence without affecting other components. Similarly, the authorization server *may* create a customized interface for managing authorized clients, scopes, and permissions without broadcasting its development complexity, structure and patterns, and technological compliance with other components. However, the constituting components *must* follow a common standard for data storage and transmission.

Expert Evaluation of the Open-Source Framework

As part of the nonempirical evaluation, we conducted an expert review. A panel of 17 experts (developers and designers) was invited to review the system and its components. We invited experts from the field, all of whom worked in the information technology industry. The review team was presented the RA and an open-source framework. An interview followed the review to determine their reaction toward the open-source framework and its components. We chose full-stack developers (7/17, 41%), front end developers (3/17, 18%), back-end developers (5/17, 29%), and system architects (2/17, 12%) with >5 years of industrial experience. The evaluation aimed to inspect the open-source framework's modifiability, extendibility, scalability, security in authentication, reusability, and code readability. The reviewers were asked to rate the evidence of these software qualities in the presented open-source framework.

The results of the expert evaluation are presented in Table 3. As shown in Table 3, experts evaluated the open-source framework as possessing most of the abovementioned capabilities. In addition to these questions, we asked open-ended

questions regarding feedback, reviews, and improvements. This feedback and reviews were considered for enhancement of the open-source framework.

Table 3. Results of expert evaluation (N=17).

Questions	Participants, n (%)					Values, mean (SD)
	1	2	3	4	5	
Component modifiability	0 (0)	0 (0)	1 (6)	8 (47)	8 (47)	4.412 (0.599)
Framework extendibility	0 (0)	0 (0)	4 (24)	5 (29)	8 (47)	4.235 (0.807)
System scalability	0 (0)	0 (0)	2 (12)	4 (24)	11 (65)	4.529 (0.696)
Security in authentication	0 (0)	0 (0)	3 (18)	9 (53)	5 (29)	4.112 (0.676)
Component reusability	0 (0)	0 (0)	2 (12)	9 (53)	6 (35)	4.423 (0.644)
Code readability	0 (0)	0 (0)	2 (12)	11 (65)	4 (24)	4.118 (0.582)

Empirical Evaluation: Case Study

In addition to nonempirical evaluation, we evaluated the framework with a small group of participants for the feasibility study for the attention-deficit/hyperactivity disorder cases in the INTROMAT project. Domain experts created a web-based intervention, and the participants were asked to interact with the intervention. The feasibility study and results are under review for publication as RCTs [49]. The feasibility study results show that the intervention system built on the top of the proposed RA can adapt to interventions, such as reminder or alert adaptation and content adaptation.

Implication of the RA

One of the essential questions is *why the RA is essential*. Our literature review revealed a lack of standard documentation, framework, and clinical guidelines on how the IDPT system should be developed [4]. As a result, developers and researchers reinvent their own version of the IDPT system, making it more complicated, less interoperable, and lacking a common foundation. Defining RA is a well-recognized method of addressing these challenges. Martinez et al [50] mentioned that RA increases development speed, reduces operational expenses, and improves software system quality. Similarly, several other

studies [6,51] have outlined the benefits of RA as it provides a template solution for a specific domain.

In the health care context, researchers, developers, and industrial partners have published RA. Therefore, one might argue *why the proposed RA is better and how it solves the identified problems*. To the best of our understanding, no RA has been reported in the psychological domain. Some related, published RAs have been compared in related work (see the *Methods* section). Furthermore, we provide a detailed architecture of the intervention system, which is part of the RA. The intervention system allows for the creation and design of interventions that can be used in several cases. Hence, both researchers and software developers can use an open-source framework or extend the framework to match their use cases. Angelov et al [52] presented a detailed framework for the analysis and design of RA. To reduce threats to validity, we used this framework [52] to analyze the proposed RA and create the contextual Table 4. We identified two problems in the current IDPT systems: (1) they lack adaptiveness, and (2) they are complex and less interoperable because of the lack of open-source standards. The *How We Propose to Increase Adaptivity in IDPT Systems* section outlines several scenarios of how the proposed RA addresses adaptiveness. We made both the RA and intervention system an open-source framework to attract researchers and developers to use it rather than reinvent it from scratch.

Table 4. Analysis of the proposed RA^a with respect to the framework presented by Angelov et al [52].

Category and questions	Details
Context	
Where will it be used?	Health care providers, hospitals, and health clinics that provide digital intervention
Who defines it?	Collaboration between psychological domain experts, software engineers, IT ^b industry partners, and HCI ^c experts
When is it defined?	With a high prevalence of mental or neurological disorders in Norway and around the world, the INTROMAT ^d project aims to provide adaptive interventions for people with mental health issues
Goals	
Why is it defined?	To adapt the intervention to reduce current higher dropouts and increase user adherence
Design	
What does it describe?	Components that are working together to form a data-driven adaptive system
How is it represented?	We used a semiformal representation of RA and described each component in detail

^aRA: reference architecture.

^bIT: information technology.

^cHCI: human-computer interaction.

^dINTROMAT: Introducing Mental Health through Adaptive Technology.

Future Work

A promising objective of this open-source framework is to adapt interventions based on user needs and preferences. However, the RA requires continuous evolution and refactoring, and so does our open-source framework for the intervention system. Our immediate future work involves (1) usability and performance evaluation of the CMS for therapists and patients' UI, (2) evaluation of adaptive strategies and their implications on user adherence using effective RCT methods, (3) maintenance and evaluation of UIs using UI experts, and (4) building and supporting analytics server end points for adaptation. There are several potential research directions for future research with this open-source framework, including the automatic structuring of modules and tasks inside a case; taxonomic or ontology-based adaptation; interoperability; and support for better user interactions, such as adaptive conversational agents.

Conclusions

To the best of our knowledge, this is the first study to create an RA and open-source framework for an adaptive IDPT system. The proposed RA uses a user profiling model to adapt and personalize interventions based on user needs. On the basis of the proposed RA, we created an open-source framework for an adaptive IDPT system. We followed the DDD architectural style and Test-Driven Development process to create an open-source framework prototype. We evaluated it using empirical (case study) and nonempirical approaches (SAAM method, expert evaluation, and software quality matrices). This paper presents an initial study, and preliminary evaluation results show that developers and researchers can extend the proposed RA to multiple health care interventions. Our immediate future work will involve extending and evaluating the framework for usability, performance, and other adaptive capabilities.

Acknowledgments

The authors owe many thanks and gratitude to their colleagues for helping with the search, refining of articles, and cross-referencing with the relevant papers. This publication is a part of the INTROMAT (Introducing Mental Health through Adaptive Technology) project [53], funded by the Norwegian Research Council (259293/o70). INTROMAT is a research and development project in Norway that uses adaptive technologies to confront these issues.

Conflicts of Interest

None declared.

References

1. Evidence-based Psychotherapy: Where Practice and Research Meet. Washington, D.C., United States: American Psychological Association; 2006.
2. Kelders SM, Kok RN, Ossebaard HC, Van Gemert-Pijnen JE. Persuasive system design does matter: a systematic review of adherence to web-based interventions. *J Med Internet Res* 2012 Nov 14;14(6):e152 [FREE Full text] [doi: [10.2196/jmir.2104](https://doi.org/10.2196/jmir.2104)] [Medline: [23151820](https://pubmed.ncbi.nlm.nih.gov/23151820/)]

3. Mukhiya SK, Wake JD, Inal Y, Lamo Y. Adaptive systems for internet-delivered psychological treatments. *IEEE Access* 2020;8:112220-112236. [doi: [10.1109/access.2020.3002793](https://doi.org/10.1109/access.2020.3002793)]
4. Mukhiya SK, Wake JD, Inal Y, Pun KI, Lamo Y. Adaptive elements in internet-delivered psychological treatment systems: systematic review. *J Med Internet Res* 2020 Nov 27;22(11):e21066 [FREE Full text] [doi: [10.2196/21066](https://doi.org/10.2196/21066)] [Medline: [33245285](https://pubmed.ncbi.nlm.nih.gov/33245285/)]
5. Bonaccorsi A, Rossi C. Why Open Source software can succeed. *Res Policy* 2003 Jul;32(7):1243-1258. [doi: [10.1016/s0048-7333\(03\)00051-9](https://doi.org/10.1016/s0048-7333(03)00051-9)]
6. Cloutier R, Muller G, Verma D, Nilchiani R, Hole E, Bone M. The concept of reference architectures. *Syst Eng* 2010;13(1):14-27. [doi: [10.1002/sys.20129](https://doi.org/10.1002/sys.20129)]
7. Kogut B, Metiu A. Open - source software development and distributed innovation. *Oxford Rev Econ Policy* 2001;17(2):248-264. [doi: [10.1093/oxrep/17.2.248](https://doi.org/10.1093/oxrep/17.2.248)]
8. *Handbook of Cognitive-Behavioral Therapies*, Third Edition. New York, United States: Guilford Publications; 1988.
9. Kohn R, Saxena S, Levav I, Saraceno B. The treatment gap in mental health care. *Bull World Health Organ* 2004 Nov;82(11):858-866 [FREE Full text] [Medline: [15640922](https://pubmed.ncbi.nlm.nih.gov/15640922/)]
10. *Adherence to Long-term Therapies Evidence for Action*. Geneva: World Health Organization; 2003.
11. Wallin EE, Mattsson S, Olsson EM. The preference for internet-based psychological interventions by individuals without past or current use of mental health treatment delivered online: a survey study with mixed-methods analysis. *JMIR Ment Health* 2016 Jun 14;3(2):e25 [FREE Full text] [doi: [10.2196/mental.5324](https://doi.org/10.2196/mental.5324)] [Medline: [27302200](https://pubmed.ncbi.nlm.nih.gov/27302200/)]
12. Portnoy DB, Scott-Sheldon LA, Johnson BT, Carey MP. Computer-delivered interventions for health promotion and behavioral risk reduction: a meta-analysis of 75 randomized controlled trials, 1988-2007. *Prev Med* 2008 Jul;47(1):3-16 [FREE Full text] [doi: [10.1016/j.ypmed.2008.02.014](https://doi.org/10.1016/j.ypmed.2008.02.014)] [Medline: [18403003](https://pubmed.ncbi.nlm.nih.gov/18403003/)]
13. Mukhiya SK, Rabbi F, Pun KI, Lamo Y. An architectural design for self-reporting e-health systems. In: Proceedings of the 2019 IEEE/ACM 1st International Workshop on Software Engineering for Healthcare (SEH). 2019 Presented at: 2019 IEEE/ACM 1st International Workshop on Software Engineering for Healthcare (SEH); May 27, 2019; Montreal, QC, Canada. [doi: [10.1109/seh.2019.00008](https://doi.org/10.1109/seh.2019.00008)]
14. Gary FA. Stigma: barrier to mental health care among ethnic minorities. *Issues Ment Health Nurs* 2005 Dec 09;26(10):979-999. [doi: [10.1080/01612840500280638](https://doi.org/10.1080/01612840500280638)] [Medline: [16283995](https://pubmed.ncbi.nlm.nih.gov/16283995/)]
15. A reference architecture for personalized and self-adaptive e-health apps. In: *Software Architecture*. Cham: Springer; 2020.
16. On online banking authentication for all: a comparison of BankID login efficiency using smartphones versus code generators. In: *Universal Access in Human-Computer Interaction*. Cham: Springer; 2019.
17. Connected health reference architecture. WSO2. URL: <https://wso2.com/whitepapers/connected-health-referencearchitecture/> [accessed 2022-04-24]
18. Wartena F, Muskens J, Schmitt L, Petković M. Continua: the reference architecture of a personal telehealth ecosystem. In: Proceedings of the 12th IEEE International Conference on e-Health Networking, Applications and Services. 2010 Presented at: The 12th IEEE International Conference on e-Health Networking, Applications and Services; Jul 1-3, 2010; Lyon, France. [doi: [10.1109/health.2010.5556588](https://doi.org/10.1109/health.2010.5556588)]
19. Rodriguez LM, Ampatzoglou A, Avgeriou P, Nakagawa EY. A reference architecture for healthcare supportive home systems. In: Proceedings of the 2015 IEEE 28th International Symposium on Computer-Based Medical Systems. 2015 Presented at: 2015 IEEE 28th International Symposium on Computer-Based Medical Systems; Jun 22-25, 2015; Sao Carlos, Brazil. [doi: [10.1109/cbms.2015.39](https://doi.org/10.1109/cbms.2015.39)]
20. universAAL – An Open and Consolidated AAL Platform. In: *Ambient Assisted Living*. Berlin, Heidelberg: Springer; 2011.
21. User models for adaptive hypermedia and adaptive educational systems. In: *The Adaptive Web*. Berlin, Heidelberg: Springer; 2007.
22. Generic user modeling systems. In: *The Adaptive Web*. Berlin, Heidelberg: Springer; 2007.
23. Adaptive information for consumers of healthcare. In: *The Adaptive Web*. Berlin, Heidelberg: Springer; 2007.
24. Intelligent user profiling. In: *Artificial Intelligence An International Perspective*. Berlin, Heidelberg: Springer; 2009.
25. Yogarajah A, Kenter R, Lamo Y, Kaldo V, Nordgreen T. Internet-delivered mental health treatment systems in Scandinavia - a usability evaluation. *Internet Interv* 2020 Apr;20:100314 [FREE Full text] [doi: [10.1016/j.invent.2020.100314](https://doi.org/10.1016/j.invent.2020.100314)] [Medline: [32426241](https://pubmed.ncbi.nlm.nih.gov/32426241/)]
26. Inal Y, Wake JD, Guribye F, Nordgreen T. Usability evaluations of mobile mental health technologies: systematic review. *J Med Internet Res* 2020 Jan 06;22(1):e15337 [FREE Full text] [doi: [10.2196/15337](https://doi.org/10.2196/15337)] [Medline: [31904579](https://pubmed.ncbi.nlm.nih.gov/31904579/)]
27. *Domain-Driven Design Tackling Complexity in the Heart of Software*. London, United Kingdom: Pearson Education; 2003.
28. *Test-driven Development By Example*. Boston, Massachusetts, United States: Addison-Wesley; 2000.
29. Mukhiya SK, Rabbi F, I Pun VK, Rutle A, Lamo Y. A GraphQL approach to Healthcare Information Exchange with HL7 FHIR. *Procedia Comput Sci* 2019;160:338-345. [doi: [10.1016/j.procs.2019.11.082](https://doi.org/10.1016/j.procs.2019.11.082)]
30. -. Sakimura, J. Bradley, M. B. Jones, B. Medeiros, and C. Mortimore, "Final: OpenID Connect Core 1.0 incorporating," 2020:2014. [doi: [10.1007/978-1-4302-6817-8_12](https://doi.org/10.1007/978-1-4302-6817-8_12)]
31. Mukhiya SK, Ahmed U, Rabbi F, Pun KI, Lamo Y. Adaptation of IDPT system based on patient-authored text data using NLP. In: Proceedings of the 2020 IEEE 33rd International Symposium on Computer-Based Medical Systems (CBMS).

- 2020 Presented at: 2020 IEEE 33rd International Symposium on Computer-Based Medical Systems (CBMS); Jul 28-30, 2020; Rochester, MN, USA. [doi: [10.1109/cbms49503.2020.00050](https://doi.org/10.1109/cbms49503.2020.00050)]
32. Sharma A, Choudhury M, Althoff T, Sharma A. Engagement patterns of peer-to-peer interactions on mental health platforms. In: Proceedings of the International AAAI Conference on Web and Social Media. 2020 Presented at: Proceedings of the International AAAI Conference on Web and Social Media; Jun 8–11, 2020; Online.
 33. Bremer V, Chow PI, Funk B, Thorndike FP, Ritterband LM. Developing a process for the analysis of user journeys and the prediction of dropout in digital health interventions: machine learning approach. *J Med Internet Res* 2020 Oct 28;22(10):e17738 [FREE Full text] [doi: [10.2196/17738](https://doi.org/10.2196/17738)] [Medline: [33112241](https://pubmed.ncbi.nlm.nih.gov/33112241/)]
 34. Nemesure MD, Heinz MV, Huang R, Jacobson NC. Predictive modeling of depression and anxiety using electronic health records and a novel machine learning approach with artificial intelligence. *Sci Rep* 2021 Jan 21;11(1):1980 [FREE Full text] [doi: [10.1038/s41598-021-81368-4](https://doi.org/10.1038/s41598-021-81368-4)] [Medline: [33479383](https://pubmed.ncbi.nlm.nih.gov/33479383/)]
 35. *Microservice Architecture Aligning Principles, Practices, and Culture*. Sebastopol, California, United States: O'Reilly Media; 2016.
 36. Zukerman I, Albrecht DW. Predictive statistical models for user modeling. *User Modeling User Adapt Interact* 2001;11:5-18 [FREE Full text] [doi: [10.1023/A:1011175525451](https://doi.org/10.1023/A:1011175525451)]
 37. Lieberman H, Fry C, Weitzman L. Exploring the Web with reconnaissance agents. *Commun ACM* 2001 Aug;44(8):69-75. [doi: [10.1145/381641.381661](https://doi.org/10.1145/381641.381661)]
 38. Höök K, Karlgren J, Wærn A, Dahlbäck N, Gustaf Jansson C, Karlgren K, et al. A glass box approach to adaptive hypermedia. *User Model User Adap Inter* 1996 Jul;6(2-3):157-184. [doi: [10.1007/bf00143966](https://doi.org/10.1007/bf00143966)]
 39. An D, Carr M. Learning styles theory fails to explain learning and achievement: recommendations for alternative approaches. *Personality Individual Diff* 2017 Oct;116:410-416. [doi: [10.1016/j.paid.2017.04.050](https://doi.org/10.1016/j.paid.2017.04.050)]
 40. Using responsive web design to enhance the user experience of chronic disease management portals for clinical users. In: *Delivering Superior Health and Wellness Management with IoT and Analytics*. Cham: Springer; 2020.
 41. Hofmeister C, Kruchten P, Nord RL, Obbink H, Ran A, America P. A general model of software architecture design derived from five industrial approaches. *J Sys Softw* 2007 Jan;80(1):106-126. [doi: [10.1016/j.jss.2006.05.024](https://doi.org/10.1016/j.jss.2006.05.024)]
 42. Kazman R, Bass L, Abowd G, Webb M. SAAM: a method for analyzing the properties of software architectures. In: *Proceedings of 16th International Conference on Software Engineering*. 1994 Presented at: 16th International Conference on Software Engineering; May 16-21, 1994; Sorrento, Italy. [doi: [10.1109/icse.1994.296768](https://doi.org/10.1109/icse.1994.296768)]
 43. Kazman R, Klein M, Barbacci M, Longstaff T, Lipson H, Carriere J. The architecture tradeoff analysis method. In: *Proceedings of the Fourth IEEE International Conference on Engineering of Complex Computer Systems (Cat. No.98EX193)*. 1998 Presented at: Fourth IEEE International Conference on Engineering of Complex Computer Systems (Cat. No.98EX193); Aug 14, 1998; Monterey, CA, USA. [doi: [10.1109/iceccs.1998.706657](https://doi.org/10.1109/iceccs.1998.706657)]
 44. Kobsa A, Koenemann J, Pohl W. Personalised hypermedia presentation techniques for improving online customer relationships. *Knowl Eng Rev* 2001 Dec 18;16(2):111-155. [doi: [10.1017/s0269888901000108](https://doi.org/10.1017/s0269888901000108)]
 45. -. -. - 2012:2012. [doi: [10.1109/conisof.2019.00014](https://doi.org/10.1109/conisof.2019.00014)]
 46. Key words for use in RFCs to indicate requirement levels. RFC Editor. URL: <https://www.rfc-editor.org/info/rfc2119> [accessed 2022-04-25]
 47. Recommendations for secure use of Transport Layer Security (TLS) and Datagram Transport Layer Security (DTLS). RFC Editor. 2015. URL: <https://www.rfc-editor.org/info/rfc7525> [accessed 2022-04-25]
 48. Reference model for service oriented architecture 1.0. OASIS Standard. URL: <http://docs.oasis-open.org/soa-rm/v1.0/soa-rm.html> [accessed 2022-04-25]
 49. Nordby ES, Kenter RM, Lundervold AJ, Nordgreen T. A self-guided internet-delivered intervention for adults with ADHD: a feasibility study. *Internet Interv* 2021 Sep;25:100416 [FREE Full text] [doi: [10.1016/j.invent.2021.100416](https://doi.org/10.1016/j.invent.2021.100416)] [Medline: [34401375](https://pubmed.ncbi.nlm.nih.gov/34401375/)]
 50. Benefits and drawbacks of reference architectures. In: *Software Architecture*. Berlin, Heidelberg: Springer; 2013.
 51. Towards a domain specific software architecture for scientific data distribution. SIP Commons. URL: <https://commons.esipfed.org/content/towards-domain-specific-software-architecture-scientific-data-distribution> [accessed 2022-04-25]
 52. Angelov S, Grefen P, Greefhorst D. A framework for analysis and design of software reference architectures. *Info Softw Technol* 2012 Apr;54(4):417-431. [doi: [10.1016/j.infsof.2011.11.009](https://doi.org/10.1016/j.infsof.2011.11.009)]
 53. Intromat homepage. Intromat. URL: <https://intromat.no/> [accessed 2022-05-23]

Abbreviations

- AAL:** Ambient Assisted Living
- AI:** artificial intelligence
- API:** application programming interface
- CMS:** Content Management System
- DDD:** Domain-Driven Design
- DL:** deep learning

IDPT: internet-delivered psychological treatment **INTROMAT:** Introducing Mental Health through Adaptive Technology

ISO/IEC: International Organization for Standardization/International Electrotechnical Commission

ML: machine learning

NLP: natural language processing **RA:** reference architecture

RCT: randomized controlled trial

RFC: Request for Comments

SAAM: software architecture analysis method **SOA:** service-oriented architecture

UI: user interface

Edited by A Kushniruk; submitted 08.06.21; peer-reviewed by H McCall, S Khodarahimi; comments to author 10.11.21; revised version received 09.02.22; accepted 21.03.22; published 20.06.22.

Please cite as:

Mukhiya SK, Lamo Y, Rabbi F

A Reference Architecture for Data-Driven and Adaptive Internet-Delivered Psychological Treatment Systems: Software Architecture Development and Validation Study

JMIR Hum Factors 2022;9(2):e31029

URL: <https://humanfactors.jmir.org/2022/2/e31029>

doi: [10.2196/31029](https://doi.org/10.2196/31029)

PMID: [35723905](https://pubmed.ncbi.nlm.nih.gov/35723905/)

©Suresh Kumar Mukhiya, Yngve Lamo, Fazle Rabbi. Originally published in JMIR Human Factors (<https://humanfactors.jmir.org>), 20.06.2022. This is an open-access article distributed under the terms of the Creative Commons Attribution License (<https://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Human Factors, is properly cited. The complete bibliographic information, a link to the original publication on <https://humanfactors.jmir.org>, as well as this copyright and license information must be included.

Publisher:
JMIR Publications
130 Queens Quay East.
Toronto, ON, M5A 3Y5
Phone: (+1) 416-583-2040
Email: support@jmir.org

<https://www.jmirpublications.com/>